DOT/FAA/AM-95/21

Office of Aviation Medicine
Washington, D.C. 20591

Controller Resource Management — What Can We Learn From Aircrews?

Charmine E.J. Härtel
University of Tulsa

Günther F. Härtel
Colorado State University

Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

July 1995

Final Report

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department
of Transportation

Federal Aviation
Administration

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-95/21		

4. Title and Subtitle	5. Report Date
Controller Resource Management - What Can We Learn From Aircrews?	July 1995
	6. Performing Organization Code
7. Author(s) Charmine Härtel, Ph.D. & Günther Härtel	8. Performing Organization Report No.
9. Performing Organization Name and Address University of Tulsa Department of Psychology 600 South College Ave Tulsa, Oklahoma 74104	10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFA-02-93P7913
12. Sponsoring Agency name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591.	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplemental Notes

This work was performed under contract from the Human Resources Research Division of the FAA's Civil Aeromedical Institute in Oklahoma City, Oklahoma as part of task AM-8-94-HRR159. Dr. Tom Hilton served as the Contract Officer's Technical Representative (COTR).

16. Abstract

This paper provides an overview of the scientific literature regarding Crew Resource Management (CRM). It responds to tasking from the Office of Air Traffic Program Management to conduct studies addressing the application of team training models such as CRM for air traffic operational and administrative tasks. The authors report that there is no single model for CRM per se. They formulate a model by integrating common dynamic elements found in many CRM programs. The literature reviewed points out that current CRM research: (a) is primarily focused on flight deck CRM; (b) is insufficient to establish the superiority of CRM training over other team training alternatives; (c) has not identified the critical components of CRM team training; and (d) is limited in part by weaknesses in assessment methods and outcome criteria. The authors identify alternative assessment methods and performance criteria currently being explored by behavioral scientists. Finally, the authors discuss potential CRM techniques for enhancing Air Traffic Control Specialist operational tasks.

17. Key Words	18. Distribution Statement
Cockpit/Crew Resource Management Team Training Air Traffic Control	Document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	40	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

Preparation of this article was supported by a grant to Charmine Härtel from the FAA Office of Aviation Medicine's Civil Aeromedical Institute.

We gratefully acknowledge Dave Schroeder and Tom Hilton for comments on this manuscript; Tom Hilton and Dana Broach for comments on the first draft of this paper; A. L. Viselli and Charlie Bono for arranging helpful meetings with controllers and trainers at Dallas/Fort Worth's Tower/TRACON and En Route Center; information provided by Alice Peck and Gail Williams on current ATTE instruction at the FAA Academy; Matthew Barney, Steven Sulkin and Larry Bailey for assistance in conducting interviews in August 1993 at Dallas/Fort Worth Tower/TRACON and Enroute Center, and in October 1993 at the FAA Academy; Eduardo Salas, Carolyn Prince and Randy Law for sharing FDCRM reports with us; and Matthew Barney, Jennifer Barney and Melissa Berry for assistance in literature search and retrieval.

The opinions, views, and/or findings contained in this report are the authors' and should not be construed as an official position policy or decision of Federal Aviation Administration, unless so designated by other official documentation, nor as representing the views of the University of Tulsa.

CONTENTS

Controller Resource Management: What Can We Learn From Aircews?

INTRODUCTION

This document provides an overview of the literature relating to cockpit and crew resource management (CRM) in aviation and explores applications of CRM to the air traffic control specialist (ATCS) occupation. The factors that spurred the initial interest in CRM will be examined to provide some insight into the development and current state of CRM training. The most common definitions, models, and dimensions of CRM, upon which the various research and training programs are based, are presented and examined for suitability to the ATCS environment. Empirical research to evaluate the effectiveness of CRM training is reviewed, to provide a better understanding of what is really known about CRM, and to illustrate the methods for and problems of training effectiveness evaluation. Finally, current team training for ATCSs is reviewed and some recommendations are made.

It may be possible to borrow and adapt current CRM methods and tools for use in air traffic controller resource management (ATCRM) programs. However, it is first necessary to thoroughly examine the history and development of current CRM programs to understand the rationale for program elements. This examination will also help decide which elements can be adapted and which need to be customized for ATCSs. Furthermore, understanding the development of CRM programs can be helpful in guiding development of new ATCRM programs. Note that in this paper CRM will be used to mean crew resource management in general, and the acronyms ATCRM and FDCRM will be used to distinguish air traffic controller resource management from flight deck crew resource management, when confusion might occur.

Crew Resource Management

Flight safety is of primary importance to all segments of the aviation community. Dramatic improvements in flight safety have been made in the last few decades. Over the years, modern aircraft have become extremely reliable, especially since the introduction of the jet engine in the 1950s, advanced cockpit technology, and redundant systems (Helmreich & Foushee, 1993, p. 5). According to Carroll and Taggart (1987), the use of flight simulators has also greatly reduced training accidents. Novice pilots can build proficiency without risking an unrecoverable error. Also maneuvers too risky to practice aloft can be simulated with safety, thereby improving pilots' technical skills for handling emergency procedures (Carroll & Taggart, 1987, p. 40). There has been a significant reduction in airline fatalities that corresponds to increased use of flight simulators. The rate of fatality (per 100 million passenger miles) declined from 1.97 in 1950, to 0.78 in 1960, 0.16 in 1970, and 0.08 in 1975, but back to 0.09 in 1980 (Yamamori, 1987, pp. 75, 81).

By the late 1970s, it became increasingly apparent, that, because mechanical and technical problems had declined, while human errors remained constant over time, human errors became the dominant cause of accidents (Cooper, White, & Lauber, 1980; Murphy, 1980). In fact, estimates of the proportion of accidents and incidents attributable to human error range from 60 to 80 percent (FAA, 1989, sec. 4a; Helmreich & Foushee, 1993. p. 5; Orlady, 1993, p. 450). Some indications of the human factors involved came from structured interviews with pilots. Their common concern was over the lack of skills training in decision making, command, leadership, and communication (Lauber, 1987, p. 7). These concerns were echoed in the now-classic National Aeronautics and Space Administration (NASA) simulation study, which was originally designed to investigate the effects of workload on crew performance (Ruffell Smith, 1979; cf. Lauber, 1987; Prince, Cannon-Bowers, Owens & Morgan, 1988, p. 3). An unexpected range in the crews' performance levels was observed. A large part of the variability could be explained by the crews' effectiveness in managing the cockpit resources (Lauber, 1987, p. 7). This led to further studies aimed at identifying such important skills as effective intra-

cockpit communication for managing the cockpit and crew (Foushee & Manos, 1981). The idea of applying classical business management concepts to training development also has its roots in these studies. The first attempts to develop training programs to improve cockpit management began at United Airlines (UAL) and KLM Royal Dutch Airlines in the late 1970s. Both programs were based on business manager training courses and concepts (Prince et al., 1988, p. 3). A NASA/industry conference was held in 1979 involving aviation management representatives from around the world (Helmreich & Foushee, 1993, p. 28). Thereafter, a number of training programs were developed and further research was pursued by commercial and military aviation interests.

Definition of Crew Resource Management

Before any training program can be developed or described, it is first necessary to define what is meant by crew resource management (CRM), and to present some of the conceptual models of how CRM is expected to improve crew performance. It should be noted that CRM originally referred to "cockpit resource management," but later was renamed to "crew resource management" to reflect involvement of crewmembers working in the cabin, dispatch, and air traffic control.

Several definitions of CRM, ranging from specific to general, appear in the literature. The definition of CRM included in the 1991 FAA Advisory Circular on CRM seems to contain the most agreed-upon elements:

> "...crew resource management [is] the effective utilization of resources available to the aircrew—information, equipment and personnel [themselves and others]—expressed in terms of individual and collective cognitive, interpersonal and motor skills; knowledge and attitudes. The scope of safe operations must include a realistic set of conditions that include the external environment (weather, ATC, dispatch, weather service, navigationaids, etc.) as well as aircraft abnormalities and emergencies." (Orlady, 1993, p. 464, quoting FAA draft).

Helmreich (1987, p. 15) defines CRM in a very similar fashion, as "the effective coordination and utilization of all available resources in the service of the flight. These resources are both inside and outside the aircraft and are both material and human, including especially the knowledge, judgment and decision-making skills of all crewmembers." Another definition of CRM states: "Cockpit resource management is the effective utilization of all available resources — hardware, software, and liveware — to achieve safe, efficient flight operations" (Lauber, 1987, p. 10; cf. Helmreich & Foushee, 1993, p. 4; Lauber, 1984). This definition borrows the nomenclature from Edwards' (1972) SHEL concept, where SHEL is an acronym for Software, Hardware, Environment, Liveware (Hawkins, 1993, p. 22).

A more general definition of CRM is put forward by Helmreich and Foushee (1993, p. 3): "CRM is thus the application of human factors in the aviation system." Contemporary human factors involves engineering, physiology, psychology, education, and the social sciences and concerns itself with all facets of the machine-human interface (e.g., Hawkins, 1993; Jensen, 1989), whereas CRM has tended to concentrate on the personal and interpersonal aspects.

Orasanu's (1990) definition of CRM considers a general context, so it can apply to non-aviation crews. It defines CRM as: "the management of information, cognitive work, communication, and actions needed to accomplish a time-constrained task" (cited in Sherman & Helmreich, 1993, p. 4). This definition, however, is so general that it covers activities that are not traditionally part of CRM. Moreover, much of CRM deals with situations that are not extremely time-constrained.

Models of Crew Resource Management

A number of models of varying complexity have been used to help explain the dynamics and the variables involved in aircrew coordination.

The SHEL model (Edwards, 1972; Hawkins, 1993, p. 22) closely matches Helmreich's (1987) definition of CRM and can be used to analyze the human factors involved in aircrew coordination. As indicated previously, SHEL is an acronym for Software, Hardware, Environment, and Liveware. The SHEL model is applied to any specific human factors setting by specifying these four components. For aircrew coordination, the four components of SHEL can be specified as follows: software includes "aircraft operating manuals,

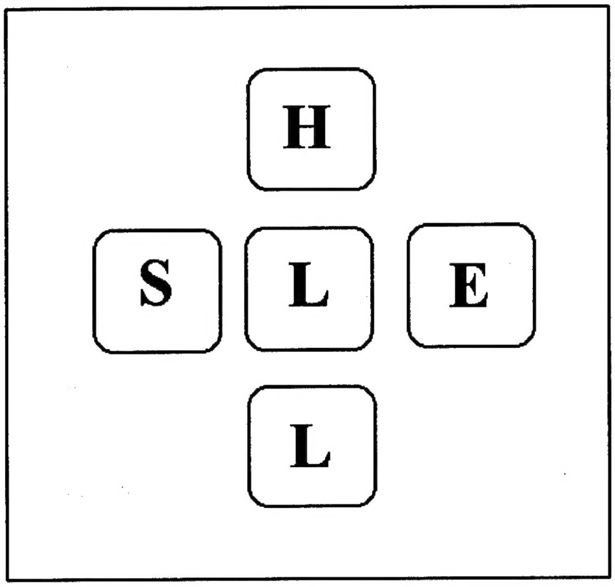

Figure 1. SHEL model.

operations bulletins, charts, and other 'information,'" and also programming, in the case of advanced technology cockpits; hardware includes "autopilots, autothrottles, and other advanced avionics;" environment includes the actual physical conditions of the cockpit, such as temperature, noise, etc.; and liveware "refers to other accessible people in the system" (Lauber, 1987, p. 9; cf. Hawkins, 1993, p. 25).

A simple "building block" diagram illustrates this model (see Figure 1). The center building block represents the individual. The lower 'L' block represents other liveware, that is, other accessible people in the system. Each building block represents an aspect of the human factors whose interaction with the center building block must be considered when evaluating a system. For example, the L-H interaction represents the interaction of the individual with the hardware, which is influenced by the design and layout of the cockpit controls and instruments, or the technical skills of the pilot. The L-S interaction includes such "non-physical" cockpit resources as the standard operating procedures (SOPs) and the information contained in "aircraft operating manuals, operations bulletins, charts, and other 'information'" necessary for flight-deck operations. Finally, the L-L interaction represents the team-member interactions, including crew communication and coordination, and so is often the primary focus of CRM interventions. Of course, a complete CRM program also addresses other aspects, as Helmreich and Foushee (1993, p. 4) put it, "CRM includes optimizing not only the person-machine interface [the L-H link] and the acquisition of timely, appropriate information [the L-S link], but also interpersonal activities [the L-L link], including leadership, effective team formation and maintenance, problem-solving, decision-making, and maintaining situational awareness." Each of the interactions is analyzed on a much more detailed level in Hawkins (1993).

Helmreich and Foushee (1993, p. 8) formulated a "three-factor model of the deter-minants of group performance" of flightcrews. The "three major components of group behavior" are: input factors, group process factors, and outcome factors (see Figure 2).

The "underlying assumption of the model is that input factors provide both the framework and determine the nature of group processes that lead, in turn, to the various outcomes." Furthermore, the outcomes affect both the input and process factors, thus resulting in a "feedback loop" (Helmreich & Foushee, 1993, p. 8). For example, Hackman (1987, p. 26) believes that experiences, both good and bad, gained during the life of the team can affect members' attitudes about teamwork, which leaves behind "a residue each time a crew terminates," and in this way, input variables "for a given organization evolve over time."

In aviation, the most important outcome factors are "safety, efficient completion of missions, [and] compliance with organizational and regulatory requirements." Input factors include: "individual aptitudes, physical condition, crew composition, organizational factors, regulatory environment, and operating environment." These are not directly changed very easily. So, interventions are aimed at the process variables, which include crew formation and management, aircraft flight control, communication skills, decision processes, situational awareness, and operating procedures. The feedback loop in this model implies that, over time, successful training of the process variables can affect input variables, resulting in cumulative improvements. A longitudinal study, which compared trainees' performance over four

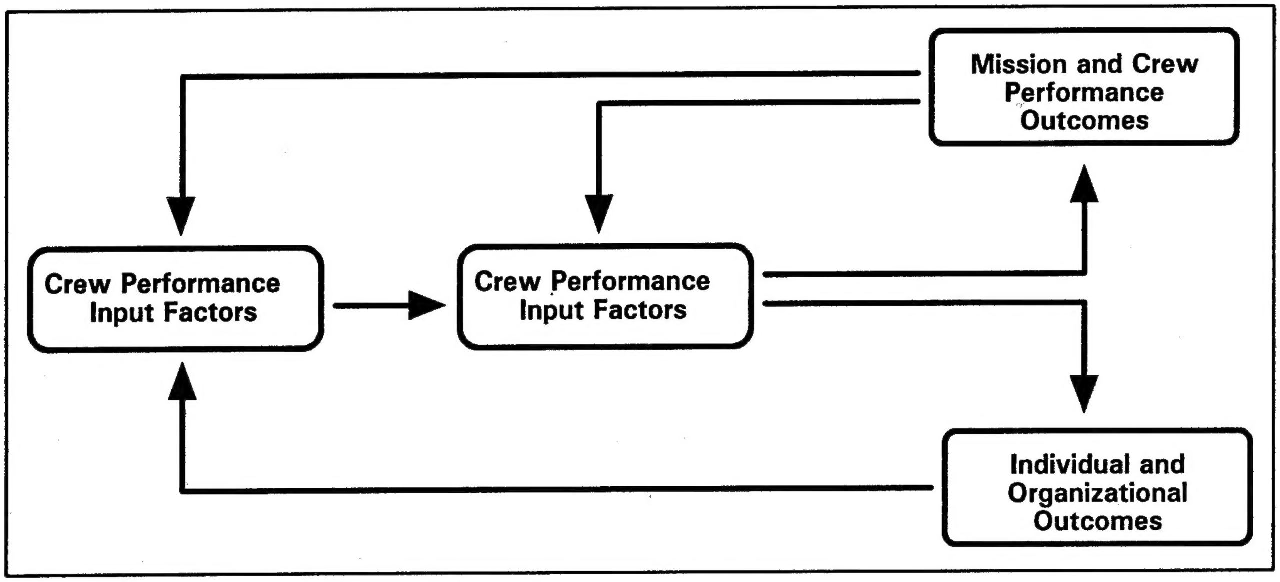

Figure 2. Flightcrew Performance Model, from Helmreich & Foushee, 1993, p. 8.

successive years, found a positive trend in test scores, thus suggesting there might be cumulative improvements in some measures of performance (Helmreich & Foushee, 1993, p. 34). This study is discussed further in a later section.

Skill Dimensions of CRM

The initial development of CRM concepts resulted from analyses of airline accident reports, flight simulator transcripts, and interviews with active pilots (e.g., Cooper, White, & Lauber, 1980; Lauber, 1987; Murphy, 1980; Ruffell Smith, 1979). From these analyses, a number of skill dimensions were identified that were considered to be important to aircrew performance. Some of these skills involved "technical proficiency and competence" of crewmembers, but others were related to "resource management or crew coordination" (Helmreich 1987, p. 15). In one sense, the technical skills are more important because, as Helmreich and Foushee (1993, p. 19) argue, "optimal team interactions and decision-making will be of little value if the crew cannot also integrate them with technical execution of maneuvers and procedures." Training of the technical skills has been the main focus of aircrew training (FAA, 1989). However, the recognition that pilot errors were "more likely to reflect failures in team communication and coordination than deficiencies in 'stick-and-rudder' proficiency," (Helmreich & Foushee, 1993, p. 7) has placed increased emphasis on crew coordination issues.

The skills that were identified as part of cockpit and crew resource management "included workload management and task delegation, situational awareness, leadership, use of available resources (including other crewmembers, manuals, air traffic control), interpersonal communications (including unwillingness of junior crewmembers to speak up in critical situations), and the process of building and maintaining an effective team relationship on the flightdeck" (Helmreich & Foushee, 1993, p. 7). Leadership skills, required primarily, but not exclusively, of the captain, include the ability to delegate tasks, provide feedback, promote crew motivation, ensure crewmembers are clear about the role expectations, and set a tone of openness that allows crewmembers to contribute their input without fear of criticism (cf. Prince, Chidester, Cannon-Bowers, & Bowers, 1992). A similar set of CRM dimensions was identified by Lauber (1979, 1987, p. 10) and included: (a) delegation of tasks and assignment of responsibilities (workload distribution), (b) establishment of priorities, (c) monitoring and cross-checking, (d) use of information, (e) problem assessment and distraction management, the avoidance of preoccupation, (f) communication, and (g) leadership.

United Airlines implemented their CRM program, using the following five CRM elements: "inquiry, advocacy, conflict resolution, critique, and decision making" (Carroll & Taggart, 1987, p. 42). Prince and Salas (1989) include the following seven skill dimensions in their CRM program: communication, situational awareness, decision-making, mission analysis, leadership, adaptation, and assertiveness (cf. Härtel, Smith & Prince, 1991; Prince et al., 1992, p. 18; Prince et al., 1988, p. 11). These skills were derived from reviews of existing CRM programs, the team training literature, and managerial assessment center findings, combined with analysis of the needs of military flightcrews (Prince et al., 1992, p. 18), and an analysis of military mishap reports (Härtel, Smith, & Prince, 1991). Although developed independently, this list of skills bears a remarkable similarity with the list of skills the FAA recommends, which are: communication, situational awareness, decision-making, team management, stress management, team review (mission analysis), and interpersonal skills (FAA, 1989; Prince et al., 1992, p. 18).

At first, the different categorizations of skills may appear to be very diverse. However, there is considerable overlap in the definitions of the skill dimensions. For example, one definition of the communication dimension may include not only technical aspects of communication, such as summarizing, reflecting, directing, and proposing, but also skills like assertiveness and problem solving (Prince et al., 1992, p. 17), while another may make assertiveness and problem solving separate categories. The skill dimensions are also interdependent, so that discussion of one dimension often involves skills from the other dimensions. For example, Orasanu (1991) defines decision making as including these components: "situation assessment," "metacognitive processes in forming action plans," "shared mental models based on intra-crew communication," and "resource management that encompasses task prioritization and delegation of specific responsibilities." Table 1 summarizes the CRM dimensions, placing the most similar dimensions on the same row. Keep in mind that the definitions of the CRM dimensions vary among researchers.

The seven skills identified in the FAA Advisory Circular (1989) may represent a developing consensus of core CRM skills (Prince et al., 1992, p. 18). These dimensions, together with the models of team functioning, provide a framework for

Prince & Salas, 1989	**FAA 1989**	**Helmreich & Foushee, 1993**	**Lauber, 1979, 1987**
situational awareness	situational awareness	situational awareness	
leadership	interpersonal skills, team management	leadership, workload management and task delegation	leadership, delegation of tasks and assignment of responsibilities, establishment of priorities
adaptation		use of available resources	use of information
communication	communication	communication	communication
mission analysis	team review or mission analysis		monitoring and cross-checking
decision making	decision making		problem assessment
	stress management	building and maintaining team	distraction management

Table 1. Comparison of CRM dimensions.

studying and understanding team interactions and human factors in the cockpit. They also help visualize how interventions might affect the crew and how their effectiveness could be evaluated. The CRM models and dimensions are presented in general forms and can be applied to different contexts by specifying the factors, variables, and interactions of the models. To do this, a better description is necessary of teams, and aircrew teams, in particular.

What is a team?

The focus of much of the CRM literature is on the functioning and performance of teams. However, attempts to define what constitutes a team are by no means universally agreed upon. Nonetheless, the various definitions of a team offer useful perspectives and they do share common elements. It is desirable to find a definition that is general enough to allow concepts based on it to be applied to many different situations, yet specific enough to exclude groups for which the concepts do not apply. Furthermore, if a suitable definition of *team* can be found, then it may be possible to draw upon elements of group and team research to guide development of CRM concepts.

It seems that the original meaning of the word *team* referred to two or more work-animals hitched together to pull a single load (American Heritage Dictionary). This definition provides a good metaphor for a team of humans. It clearly distinguishes a team of animals from any other group or herd of animals. It illustrates the coordination and interdependence that is characteristic of teams. Team members are, in a sense, "hitched" together to perform a single task.

The various definitions of teams in the literature seem to share these elements but provide more specific details. For example, Hackman and Helmreich (1988, p. 285) define a team by the following three criteria: (a) the team must be a real group, that is, the team must be an "intact social system complete with boundaries and differentiated roles among members," (b) the team must have "one or more tasks to perform, resulting in potentially measurable outcomes of members' collective work," and (c) the team must be a "group that operates within an organizational context" (Hackman & Helmreich, 1988, p. 285).

A different wording of this definition is given by Hackman (1986), who defines a team as "a small social system in which (a) membership is clearly defined (i.e., one can readily distinguish members from nonmembers), (b) members have differentiated roles to play in pursuit of some common purpose, and (c) the team as a whole manages transactions with other individuals and groups as it goes about its work (adapted from Alderfer, 1977)." Both of these definitions mention an organizational context, leading to three levels of analysis: individual, team, and organizational.

Finally, coming from a military perspective, Salas et al. (1992, p. 4), define a team "as a distinguishable set of two or more people who interact, dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission, who have each been assigned specific roles or functions to perform, and who have a limited life span of membership." Furthermore, they argue that types of teams "fall on a continuum" ranging from "highly structured, interdependent teams" to "teams whose members interact minimally and perform tasks in a group context." Aircrews would tend to fall near the "highly structured, interdependent teams" end of the continuum.

The aircrew team

The above discussion highlights the complexity associated with defining a team. However, when examining common elements representing different perspectives, it seems to us that a good working definition for an aircrew team pivots on the notion of interdependence advocated by Salas, et. al. (1992). We define an aircrew team as being: *Two or more individuals, each with specified sociotechnical roles, who work interdependently within a proscribed context toward a common mission objective.* This definition sets a minimum size for a team (two), acknowledges the existence of role boundaries tied both to interpersonal and technical factors (e.g., pilot-in-command, engineer), in a context that is identified by distinguishing and finite attributes (e.g., time, space), and that involves interdependent behavior.

Helmreich and Foushee (1993, p. 20) note that: "Flightcrews are teams with a designated leader and clear lines of authority and responsibility."

The typical airline flightcrew consists of two or three members: the pilot (or captain), the copilot (or first officer), and, in some cases, a flight engineer/second officer (Hackman & Helmreich, 1988, p. 286). Most modern commercial aircraft have been certified for two member crews (Blanchard, 1994).

On the flight deck, each crewmember has a definite rank associated with the experience and certification in the model of aircraft flown which imparts authority and responsibility assigned to their assigned role. A newly-hired pilot usually starts as first or second officer (depending on flightcrew requirements), and upon reaching seniority, qualified crewmembers are promoted when a vacancy occurs until they reach the rank of captain[1]. Duties and responsibilities are clearly defined for each crewmember. The captain has ultimate responsibility for the safety of the flight and even airline management cannot order the captain to take a flight that he or she considers to be unsafe. The first officer shares flying duties with the captain, but the captain can take over at any time. The first officer, however, can take control from the captain in the event that the captain becomes incapacitated. The second officer controls and monitors the mechanical systems (Hackman & Helmreich, 1988, pp. 286,287).

Each flight deck crewmember's activities are clearly defined for every phase of a normal flight and even many non-normal situations. During normal flights, the crew's workload is predictable, but very uneven. There is a substantial workload near the beginning of a flight and then again near the end, with very little work in between. Occasionally, abnormal incidents punctuate the "long periods of routine activity" with "demands for intense and highly interdependent teamwork" (Hackman & Helmreich, 1988, p. 288). Aviators often characterize their work as "hours of intense boredom punctuated by moments of sheer terror."

The short time aircrew teams spend working together is another distinguishing characteristic of flight deck teams. Often, airline crews work together for one month, and are "composed of individuals who typically did not choose to work together" (Hackman & Helmreich, 1988, p. 288). Team formation happens rapidly, since crewmembers "have little time to get to know one another before their first period of demanding collaborative work begins" (Hackman & Helmreich, 1988, p. 288). The length of time that a team has been together can affect team performance. A study on the effects of fatigue on performance found, surprisingly, that "fatigued crews (who had just flown together) were rated as performing significantly better and made fewer serious operational errors than the rested, pre-duty crews" who had not recently flown together (Foushee, Lauber, Baetge, & Acomb, 1986). The positive effects of having worked together as a team outweighed the negative effects of fatigue.

The design of teams' task and work area can also affect how one might define a team. The introduction of advanced technology, or "glass cockpit" aircraft illustrates how automation can affect crew interactions, in particular, crew communications. Wiener (1993, p. 213) found some unexpected results when he compared two advanced technology cockpits to their respective predecessors. He found "that during the periods of high workload, climb and descent, the 737-300 (glass cockpit) pilots were more communicative than those in the 737-200 (standard cockpit)." But he found the opposite result when comparing a glass cockpit 757 with a standard cockpit 757. He noted that very little communication seemed to go on between crewmembers of the glass cockpit 757. During one phase of the flight, the only intra-cockpit communication was "between each pilot and his CDU (control-display unit)" (Wiener, 1993, p. 213). It seems that the effect of automation is not uniform, but depends on the specifics of the implementation. CRM researchers note the need to learn how "best to train crewmembers to interact most effectively with 'electronic crewmembers'" (Helmreich & Foushee, 1993, p. 39).

Factors affecting the composition and culture of the team may affect how it is defined. Flight deck researchers have noted that the typical airline crew continues to evolve as a team in response to several factors. Economic issues and improving technology,

[1] An interesting exception to the seniority ranking occurs in some airlines that offer senior captains over the age of 60 the option of continuing to fly as flight engineers (Blanchard, 1994).

such as automation and more compactly designed cockpits (Komons, 1987) have made it possible to eliminate the role of the flight engineer in many commercial aircraft. There is evidence that communication patterns in a two-person group are very different from those in a three-person group (Wiener, 1993). Also, eliminating the flight engineer position means that the entry-level position in the cockpit is now the co-pilot, so the pilot flying may have fewer hours in the air than was typical of three-person crews. Another factor that may result in less experienced (in terms of flight hours) airline pilots is the emerging trend of some air carriers to sponsor ab initio (from the beginning) pilot training.[2] Ab initio training has been motivated in part by the declining supply of military trained pilots (Orlady, 1993, p. 455; Wiener, 1993, p. 201). It is likely that, as a larger proportion of pilots have non-military backgrounds, the prevailing cockpit culture may also change. Atypical airline crews are also found in "ultra-longhaul flight operations," such as international flights (Helmreich, 1993, p. 3), which "require augmented crews," i.e., backups for the regular crew. Helmreich (1993) speculates that "such extended teams raise issues of leadership, shift changes, and the utilization of extra crewmembers in emergency situations," which would affect CRM and CRM training.

Mission differences are becoming increasingly identified as a potential factor in team definition. Cockpit layouts and team configurations are much more variable in military flying than in passenger and cargo flying. Unlike typical air carrier crews, military crews can range in size from a single pilot to more than a dozen individuals. Military aircrews often fly in groups of planes, and thus must coordinate as closely among aircraft as within the aircraft. Lines of authority are more complex. The captain of an airliner is the single individual in charge, but a military flight may have two commanders: one who leads the mission and one who is responsible for the safety of the flight. In nearly all passenger and cargo transport flying, comfort and safety dominate over mission accomplishment (on-time delivery at the scheduled destination).

The many factors affecting definition of a team potentially impact CRM design. Except for some Military Airlift Command (MAC) passenger flights, military flying normally weights mission accomplishment (e.g., actual or simulated combat) as equal to or of greater importance than comfort and safety. For example, bombing runs must be made both in training and in practice under adverse weather, speed, separation, and other conditions. Therefore, such mission-dominant flying is far more likely to involve risk. Many of the military-specific missions also require heightened coordination. For example, the missions may include "dogfights" between aircraft, or extremely low altitude nap of the earth (NOE) flying, where, in most cases, the crew has less than half a minute from the first sign of a problem until impact (Prince & Salas, 1993, pp. 341-344). Leedom (1990) concluded that the time pressure for Army rotary wing crews, along with other differences in the type of flying required by the Army and the civilian aviation communities, makes implementation of much that is taught in current CRM programs questionable.

How is CRM taught?

A review and analysis of current CRM training methods, principles, and programs can be useful for understanding how best to implement CRM training for ATCSs. CRM training has evolved since its inception, with some authors referring to current methods as "third generation CRM" training (Helmreich & Foushee, 1993, p. 30). A variety of CRM programs abound in the industry, each with some special emphases and philosophies. There are commonalities in the development and implementation of the various programs.

The development of CRM programs tends to include the following three steps in one form or another (see Figure 3). First, the objectives of the CRM program are defined, based on some form of needs analysis (Gregorich & Wilhelm, 1993, p. 190; Prince et al., 1992, p. 18). This can include

[2] Currently, no major U.S. carriers are sponsoring ab initio training, however, several foreign carriers use ab initio programs and the background of pilots in some U.S. commuter airlines might be considered ab initio.

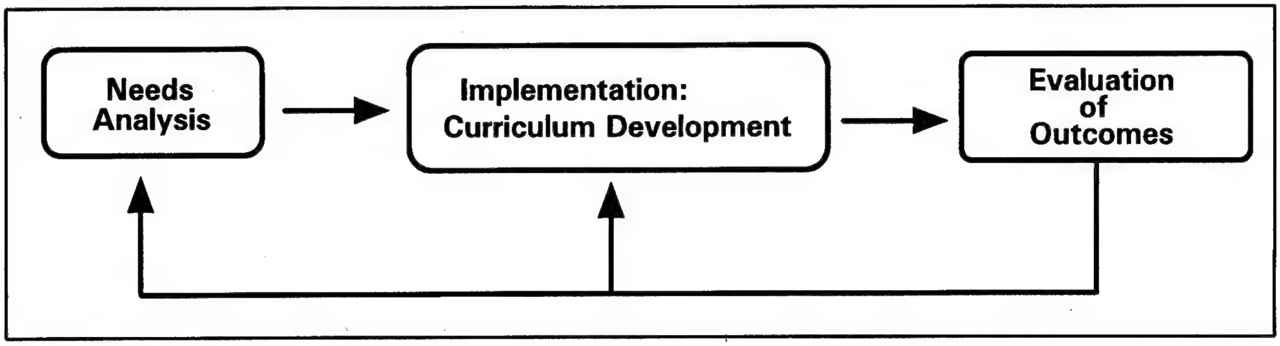

Figure 3. CRM training program development cycle.

analyses of accident and error reports, observations from flight simulator sessions, and questionnaires and interviews to assess attitudes and skills of crewmembers (e.g., Cooper, White, & Lauber, 1980; Gregorich & Wilhelm, 1993, p. 191; Lauber, 1987; Murphy, 1980; Ruffell Smith, 1979). Second, these objectives are then used to create a training program curriculum. How this step is accomplished depends on the developer's particular CRM philosophy, and organizational and regulatory factors. Illustrating the effects of CRM philosophies on training implementation, Prince et al. (1988, p. 6) note that "Sellards, of US Air, in a cognitive approach, has argued that training must be done in the classroom; Helmreich . . . has suggested that altering attitudes is the way to affect cooperative behavior; and Hackman . . . focuses on training team management skills. None of the three is able to offer evidence that a CRM program based on their approach has been more effective for reducing human error in the cockpit than a CRM program based on another approach." In the third step, the program is evaluated to judge its effectiveness. Evaluating effectiveness not only helps justify program expenditures, but it also provides information necessary to modify program elements to improve instructional techniques, update curriculum, and meet changing job requirements. The results of the program evaluations may lead to modifications in the curriculum and possibly a re-evaluation of training needs.

Historical Roots of CRM

CRM training programs have been traced back to team training in the military as early as 1956 (Brown, 1987; Prince et al., 1988, p. 4), although other names, such as co-pilot training, were used for these programs. In the early 1970s, following several "pilot error" accidents in the Pacific, Pan American World Airways initiated a program of "crew concept training." The first step was to establish a "flight operations review team" to review current training methods and make recommendations (Butler, 1987, p. 61; Helmreich & Foushee, 1993, p. 7). The ensuing "crew concept training" meant that "both simulator training and checking were to be conducted not as single-pilot evolutions but in the context of a full crew" (Helmreich & Foushee, 1993, p. 7). Furthermore, manuals were revised to prescribe responsibilities for team activities and communications. The program did not, however, provide any specific training in how to communicate or coordinate as a team (Helmreich & Foushee, 1993, p. 7).

Following a NASA/Industry workshop concerning human factors in aviation in 1979, a number of CRM programs began to be developed. These programs typically consisted of lectures or videotaped presentations. Videos are used to review the role of human factors in accidents with the goal of affecting attitudes about appropriate flight deck management (Helmreich & Foushee, 1993, p. 28). Some programs also included some tests and exercises to make students more aware of their management styles and their effect on other team members. For example, tests to assess subjects' concern for people and for performance were used to categorize their management style using the "management grid" (Blake & Mouton, 1982, cited in Wiener et al., pp. 402, 403). One example of an exercise involves one group of students building a cardboard structure according to directions given by a second group of students. In this way the need for communication and

coordination is demonstrated and appropriate skills practiced. Another program, United Airlines' Command, Leadership and Resource Management (CLR) program, includes a regular review and practice of CRM (or CLR) concepts as part of the required annual simulator training session, using a line-oriented flight training (LOFT) approach (to be discussed further below).

Basic Elements of CRM

From these roots arose a basic structure of a fully implemented CRM training program, which includes "three distinct phases: (1) an *awareness phase*, where CRM issues are defined and discussed; (2) a *practice and feedback phase*, where trainees gain experience with CRM techniques; and (3) a *continual reinforcement phase*, where CRM principles are addressed on a long term basis" (FAA Advisory Circular, 1989, sec. 5c; Gregorich & Wilhelm, p. 175).

Awareness Phase

The awareness phase of CRM training is usually the easiest and most inexpensive of the three phases to implement. It "involves communicating basic knowledge of human factors concepts that relate to aviation" with "a focus on the crew-level" (Helmreich & Foushee, 1993, p. 4). The FAA CRM advisory suggests that this can be accomplished using a combination of "lecture presentations, discussion groups, role-playing exercises, computer-based instruction, and videotaped examples of good and poor team behavior" (FAA, 1989). United Airline's CLR program developed a special home study textbook that is then reinforced by seminar-format classes (Carroll & Taggart, 1987, p. 42). The Air Force's CRM program uses interactive videodisc to present its phase one material (Prince & Salas, 1993, p. 353). The CRM program at People Express presents the first phase in a home-study format with a combination of workbook, audio tapes, and "self-evaluation" flash cards (Bruce & Jensen, 1987, p. 53). Trans Australia Airlines (TAA) uses a set of five video tapes that were filmed in the simulator and at the airport to realistically portray aircrews modeling good and bad CRM behaviors and to foster discussions of CRM concepts (Margerison, McGann, & Davies, 1987, p. 97).

The content of phase one usually consists of explanations and descriptions of the CRM model and dimensions. For example, the FAA circular suggests presentation of the seven CRM skills: communication, situational awareness, problem solving/decision making/judgment, team management, stress management, team review (mission analysis), and interpersonal skills (FAA, 1989). Typically, the importance of these behaviors is motivated by reviewing actual cases in which examples of such behaviors either caused or prevented an incident or accident.

Practice and Feedback Phase

The purpose of the practice and feedback phase is to help pupils apply the CRM principles that were studied in the first phase. This is important to the effectiveness of the training, since "individuals may accept, in principle, abstract ideas of [CRM concepts] . . . , but may find it difficult to translate them into concrete behaviors on the flight deck" (Helmreich & Foushee, 1993, p. 26). The FAA circular recommends that the practice of the CRM skills be provided in the simulator, and that the feedback be facilitated by using simulator session video tapes (FAA, 1989). This would preferably be accomplished by full simulation of an entire flight using a complete crew, known as line-oriented flight simulation (LOFT), or at least by some "other simulated or actual operation scenario." Crewmembers would be evaluated both for their technical expertise and use of CRM concepts as they respond to a series of incidents, which may or may not lead to emergencies (FAA, 1989). The session would be videotaped, so that the performance can be critiqued by crewmembers and by others. The use of LOFT in phase two presents a logistical problem, since phase one can be presented to a whole classroom of trainees, but simulator sessions must be scheduled for one flightcrew per simulator. The limited access to simulator time can often result in a considerable delay between phase one and phase two (Prince et al., 1992, p. 14). Furthermore, the use of LOFT does not allow feedback to be given immediately following "an action or event, and there is no evidence to determine whether timely feedback or realism are more valuable for training" (Prince et al., 1992, p. 22). These reasons may have contributed

to the fact that, as Helmreich and Foushee (1993, p. 29) observed, "although there has been a great proliferation of CRM courses, there has not been a parallel growth in the use of CRM/LOFT to provide practice and reinforcement."

One alternate strategy to LOFT in the feedback phase is to use artificial, complicated group tasks and exercises that require teamwork and provide opportunities to apply CRM concepts (FAA, 1989). Another approach is to provide "behavioral modeling," and "active practice with feedback" (Prince et al., 1992, p. 27). For example, United Airlines uses a "structured learning process that allows crewmembers to learn, firsthand, how to use" the information presented in phase 1. "Crewmembers are able to analyze how they react to various leadership styles in the cockpit and how their own behavior can affect operational outcomes." This is all done in a "seminar environment" without the use of a simulator (Carroll & Taggart, 1987, p. 42). The advantage is that crewmembers can "address this new area without the burden of the detailed attention normally necessary in the operation of the cockpit" (Carroll & Taggart, 1987, p. 42). Details of the "structured learning process" are treated as proprietary by the airline and are not made available. Curriculum details are often difficult to find, as Prince et al. (1992) note, "specific course content is not generally available for review." United also uses a role-play approach to simulation in their captain upgrade training to provide practice and feedback (Taggart, 1993, p. 16). These "simulations" use scenarios derived from actual line-operations problems. To recreate the scenario, participants, following written instructions, play the roles of the different crewmembers. The new captain then has an opportunity to solve the problem using available resources. After two captains separately go through the scenario, their performance is discussed, compared, and reviewed.

With further research, more limited simulators, such as part-task trainers, or flight training devices (FTDs) may also be useful. Orlady (1993, p. 469) comments that "the level of fidelity required for the most effective teaching and the best use of limited fidelity are both still controversial. . . . Almost inevitably, increased fidelity means increased costs. In addition, there is considerable evidence that in some cases, unneeded fidelity can actually decrease training effectiveness by causing unnecessary distractions." Prince et al. (1992, p. 28; cf. Härtel, 1991, 1994a) have experimented with low-fidelity simulations based on desktop computers and off-the-shelf software and components, and have suggested that "low fidelity simulation . . . appears to provide enough salient cues to elicit coordination behaviors," but found that there are few data to suggest which "salient cues" are necessary for CRM training, or what level of simulation may be optimal, considering both costs and training effectiveness (Prince et al., 1992, p. 28). Helmreich (1993, p. 4) concurs: "We need to know much more about the effectiveness of this type of training in less elaborate facilities ... more limited simulations that do not encompass a full mission (SPOT: Special Purpose Operational Training) in communicating and reinforcing CRM concepts."

Continual Reinforcement Phase

The purpose of recurrent training is to reinforce and instill the CRM principles to ensure they become part of crews' everyday operation and that training effects are enduring (Helmreich, 1987, p. 19). The FAA CRM circular concedes that changing attitudes and behaviors that "have developed over a crewmember's lifetime" requires repeated exposure (FAA, 1989). Data suggest that "without reinforcement the impact of CRM decays" (Helmreich & Foushee, 1993, p. 37). Furthermore, a long-term goal is for CRM to become an "inseparable part of the organization's culture" (FAA, 1989). It is hoped that integrating team training into recurrent training programs will, in time, help change organizational norms and attitudes from an individual focus to a team focus (Helmreich & Foushee, 1993).

Recurrent training usually includes a session of "refresher curriculum" (FAA, 1989), and a practice and feedback session, typically involving a videorecorded LOFT session. This session is often part of required annual and semi-annual recurrent technical simulator training and assessments. The FAA has recommended that "these recurrent CRM exercises take place with a full crew—each member operating in their normal crew position" (FAA, 1989). Some of the programs that do not provide LOFT as part of phase two, do provide it

as part of the recurrent training (Carroll & Taggart, 1987, p. 42; Prince & Salas, 1993, p. 353; Shroyer, 1987). In the United States, LOFT has been integrated in CRM programs at "United, Horizon Airlines, Delta, Continental, and units of military aviation" and "Northwest Airlines, Delta, US Air, and Comair are in the process of implementing them" (Helmreich & Foushee, 1993, p. 29).

Videotaping. Practice and feedback are thus the critical features of both phase one and two. Helmreich (1987, p. 19) emphasizes that "LOFT with videotape feedback is one of the most powerful tools we have. I [Helmreich] am convinced CRM without the chance for practice and self-observation that comes with LOFT will be relatively ineffective." Videotaping of the sessions is considered to be essential for the feedback session (Orlady, 1993, p. 461). It allows the crew to view their own performance more objectively and to develop some self-awareness of their interactions (Prince et al., 1992, p. 24). These videotapes and LOFT sessions should be used for self-evaluation and training only, and not include any type of "check or final evaluation" (Orlady, 1993, p. 461). All LOFT recordings are typically erased following self-evaluation to ensure the crew of confidentiality and make them comfortable with having their performance recorded. This also protects the organization from possible legal exposure (Helmreich et al., 1993, p. 484).

Facilitators. CRM training sessions are typically conducted by line pilots with an interest in training (Prince et al., 1992, p. 23). Instructors without flying qualifications or experience would have difficulty understanding the dynamics and challenges facing flightcrews and would also have little credibility in the eyes of experienced pilots (Prince et al., 1992, p. 23). This "represents a trade-off between high-credibility with the audience and in-depth knowledge of the psychology of group dynamics" (Prince et al., 1992, p. 23). Thus, the trainer is more of a facilitator than an instructor. Nevertheless, the selection and training of these facilitators is considered to be very important (Helmreich, 1987, p. 19). They must be "highly skilled in all areas related to CRM performance, and they should also be expert observers of crew coordination dimensions" (FAA, 1989). The importance of training in the observation of CRM skills is illustrated by one study which found that, without performance rating instruction, self-ratings by managers who had watched their videotaped performance "were no more accurate than those made by subjects who had not been exposed to their own videotaped performance" (Prince et al., 1992, p. 24). The sensitivity and ability to control the dynamics of a group are important skills for facilitators. Helmreich & Foushee (1993, p. 38) found that the "group dynamics of particular seminars" contributed to crewmembers rejecting CRM training.

LOFT

Because LOFT plays a central role in much of the CRM literature, further discussion of issues surrounding LOFT is warranted. First, a definition and description of LOFT: LOFT is the acronym for line-oriented flight training. Line-oriented implies an adherence to procedures used in day-to-day line operations. As Butler (1987, p. 63) puts it, "LOFT is a total training concept which requires line airmen to occupy the same crew position, the same duties, the same responsibilities, and the same roles they have in day-to-day operations. It covers a flight from crew report-time to the blocks at destination." In a typical LOFT session, "a complete two- or three-person crew undergoes the simulation of an entire flight between cities. The goal of the simulation is to reproduce the complete flight environment, including dispatch releases, weight and balance computations, en route weather, and communications with the cabin crew, air traffic control, and company operations. Typically, one or more abnormal or emergency situations are introduced during the flight" (Hackman & Helmreich, 1988, p. 305). Historically, the purpose of LOFT has been to provide crew training and not to evaluate or qualify individuals or teams (Orlady, p. 461; Prince et al., 1993, p. 70).

LOFT is one of three types of line-operational simulation (LOS), and should be distinguished from both special-purpose operational training (SPOT), and line-operational evaluation (LOE) (Butler, 1993, p. 233; Prince et al., 1993, p. 70). SPOT simulations are used to learn and practice specific skills or situations, and so do not require an entire flight to be simulated in complete

realism, but, instead may be composed of several legs, and permit direct instructions and interruptions. LOE, on the other hand, is used for testing and evaluating crewmembers. LOE scenarios are designed to satisfy specific testing requirements. As mentioned earlier, either LOFT or SPOT could be used for CRM training, although LOFT is generally recommended, at least until other options are better understood (FAA, 1989; Orlady, 1993, p. 469; Prince et al., 1992, p. 28).

A critical factor in the effectiveness of LOFT for CRM training is the design of the scenarios. Scenarios should be designed to be realistic and believable, but also must include situations that challenge the crews' teamwork (Butler, 1993, p. 234; Prince et al., 1993, p. 70). The problems or emergency situations should be representative of problems that might be encountered in actual flight and should not be contrived or include any "tricks designed to catch the crewmembers" (Prince et al., 1993, p. 72).

A survey of LOFT scenarios in use at several airlines revealed that "the most successful LOFT designs usually focused on relatively simple technical problems or system failures, introducing additional complexities such as communications failures or ambiguous warning lights. Time compression and ATC restrictions also contributed effectively to the more successful LOFT scenarios" (Butler, 1993, p. 250). Team coordination seems to become most important during periods of heavy workload (Lauber, p. 7, 1987). Despite publication of recommendations and regulations for LOFT use (e.g., Butler, 1993; FAA, 1990; Prince et al., 1993), the survey found that "the design of LOFT scenarios varied considerably from one airline to another. The greatest difference noted is in the level of technical skills orientation. This varied from presenting no technical problems to extremely high technical orientation" (Butler, 1993, p. 249). The biggest challenge seemed to be to create realistic decision situations. Butler (1993) observed that "most designs did not have options available, so that decision-making was relatively simple and did not encourage interaction between crewmembers [for an exception see Härtel, 1991, 1992, 1994a]. . . . Many scenarios were nothing more than a full mission simulation with a problem requiring the use of an abnormal checklist. No challenge, no decisions." Furthermore, few of the scenarios seemed to be designed to address specific CRM skills (Butler, 1993), despite recommendations that each scenario be based on specific training objectives (Lauber & Foushee, 1981; Prince et al., 1993). Even well-designed scenarios can have poor results if the facilitators are inadequately trained. In some cases, instructors failed to follow the LOFT scenario syllabus faithfully, compromised flight realism, or failed to utilize the videotapes for feedback during debriefing (Prince et al., 1992, p. 23). Another concern with LOS training and evaluation is the limited number of scenarios in use, which might become generally known to flightcrews (Prince et al., 1992, p. 22). It is unknown what the "effect of prior knowledge of a scenario's events on performance" would be (Prince et al., 1992, p. 22).

Other Occasions for Training

Once a CRM program has been established, it becomes part of the regular training regimen. Airlines typically have four types of training. The first type is for newly-hired crewmembers, including ab-initio training; the second type is for newly promoted crewmembers; the third type is for transition or conversion training to different aircraft; the fourth is recurrent training, which is mandated by the FAA and affects all pilots (Orlady, 1993, p. 455). To fully integrate CRM into an organization, appropriate CRM training should be developed for each of these types of training. For example, at United Airlines CRM is "part of upgrade and other routine training. CRM concepts are also part of annual enroute checks" (Carroll & Taggart, 1987, p. 42; cf. Taggart, 1993).

Training Philosophies

There seem to be two major approaches to CRM training. The first approach makes change in attitudes about flight deck management the training objective. The second approach makes specific skills and behaviors involved in flight deck management the training objective.

The attitude approach is based on the idea that individual pilots differ with respect to "three broad categories of characteristics," namely, ability, personality, and attitudes (Helmreich, 1987, p. 16). The ability to operate an aircraft is a

prerequisite requirement of pilots and has traditionally been the focus of training. Personality is, by definition, difficult, if not impossible to change. Therefore, the training objective is to change CRM attitudes. Proponents of the attitude approach argue that attitudes affect performance in line operations (Helmreich, 1984, 1987; Helmreich, Foushee, Benson, & Russini, 1986). To support this view, an attitude survey (CMAQ - Helmreich, 1984) was developed and used to gather evidence that training affects attitudes, and that attitudes are related to evaluations of LOFT and line performance (Helmreich et al., 1990). A database of responses to this instrument by CRM trainees has been accumulated over several years. These data, and assertions based on them, will be discussed further in subsequent sections.

The skill-based approach emphasizes skill training instead of changing attitudes. Authors such as Prince and Salas (1993) argue that one weakness of the attitude approach is that it seems to ignore the need for specific skills for good cockpit management and coordination. In contrast, the specific-behavior approach to training advocated by Prince and Salas "... seeks to identify particular behaviors within skill dimensions that result in effective aircrew performance, and thus, require training. . . . In order to train and evaluate crew coordination, the construct of coordination is distilled into knowledge, skills and abilities (KSAs) and expressed as a set of specific targeted behaviors" (Prince et al., 1992, p. 26).

Training philosophies are important because they define criteria for successful training. If attitude change is the primary goal of training, then sponsors are going to be looking for evidence that CRM training has led to attitude change. If skill development is the primary goal of training, then sponsors will be looking for evidence that CRM has changed behaviors. In either case, underlying training philosophies have associated assumptions that influence the goals that sponsors establish. For example, some training sponsors assume that attitude change results in changes in behaviors, and that in turn, workplace climate and organizational culture will be impacted. Sponsors of skill-based training may make similar assumptions with the result that comprehensive assessments often reflect very similar criteria for assessing effectiveness. Differences in philosophy become problematic, when limited assessment resources require more narrowly-focused evaluations.

Assessing CRM Program Effectiveness

Efforts to measure the performance of crewmembers seem to be part of most CRM training programs (Gregorich, Helmreich & Wilhelm, 1990, p. 682; Gregorich & Wilhelm, 1993, p. 182; Hackman & Foushee, 1993; Hackman & Helmreich, 1988; Helmreich, 1984, p. 586; Helmreich et al., 1986, 1987, 1990; Prince et al., 1988, p. 6; Prince & Salas, 1993, p. 352). Three main purposes for performance measurement can be identified: research, performance evaluation, and training.

Performance measurements for the purposes of evaluation and training require individuals to be identified, but have an important difference. "Training-oriented assessments . . . are consequential mainly for his or her own learning and development. Evaluation-oriented assessment . . . may . . . affect the size of one's raise, the probability of a promotion, or even the security of one's job" (Hackman & Helmreich, 1988, p. 305). Therefore assessments for these two purposes should be kept separate.

Research into the effectiveness of CRM training has the dual goals of trying to empirically prove that CRM training is effective in improving crew performance, and trying to find ways to improve CRM training methods and curriculum by identifying portions that are ineffective or need to be updated to meet changing needs. Performance measurements for research purposes can be done anonymously, using only group statistics or identifying individuals by a confidential code, thus avoiding some of the privacy and personal jeopardy issues associated with performance assessments.

Empirical evidence of CRM effectiveness may be important for several reasons. First, it is possible for training to have a negative effect on crew performance (Helmreich & Foushee, 1993, pp. 5, 38), actually reducing safety, so it is important to show that any training effect is positive. Secondly, evidence of effectiveness may be necessary to justify the costs of the training, which can be significant, especially when extensive use of simulation is incorporated. Third, adaptation of CRM

concepts to contexts outside the cockpit, including application to the ATCS profession, requires evidence of effectiveness. Finally, there is also an academic interest in learning more about how to effectively train teams and to validate the theories underlying CRM concepts.

Multiple levels of evidence of CRM program effectiveness may be addressed. The first level simply compares CRM training to no CRM training, or before and after training. The next level compares CRM training to competing or alternate training methods, simply to see if there are significant differences among them. Finally, cost analyses could be done to determine whether CRM training is cost-effective, which components or combinations of components of CRM training are most cost-effective, and what level of investment in CRM training is optimal. The matter of evaluating CRM components is discussed further in a subsequent section. Even the most basic level of analysis presents formidable problems, due to the scope of typical CRM programs and the difficulty of defining and measuring such training outcomes.

The first step must be to show that CRM training is indeed effective. The ultimate assessment of the effectiveness of CRM training would ideally "be measured by how successful we are in preventing accidents and incidents which stem from inadequate or ineffective cockpit resource management" (Lauber 1987, p. 5). This is difficult to do, as Helmreich and Foushee (1993, p. 9) explain: "while the most compelling measure of effectiveness in aviation would be a decrease in the frequency of accidents, such accidents are already so infrequent that reliable statistical evidence can only be found by aggregating data over extremely long periods of time." For example, what improvement could be made to the safety record of Trans Australia Airlines (TAA), which has had an "accident-free record" for its 25-year existence (Davidson, 1987, p. 88)? "Accordingly, criteria of group performance need to be drawn from surrogate measures such as records of operational errors, expert ratings of crew effectiveness, and measures of attitude and job satisfaction" (Prince et al., 1992, p. 29).

Concern for assessing CRM effectiveness led to the Aircrew Coordination Training (ACT) evaluation project undertaken by NASA Ames and led by Professor Robert Helmreich at the University of Texas. Currently, ACT is the only large scale attempt to evaluate CRM training accessible to the scientific community (Prince et al., 1992). Based on the CRM model depicted in Figure 2 on page 9, this project uses "two types of ACT effectiveness measures: outcome and process. Outcome measures include aircraft incidents, crewmember attitudes, ratings of crew performance, and indirect measures (e.g., organizational climate). Process measures involve crew interaction and behavior as measured by expert raters and communication analyses" (Prince et al., 1992, p. 30). Typically, the evaluation of a training program's effectiveness includes several assessments, including "reactions to the program materials, attitude change, and a learning of the concepts, with an emphasis on demonstration of transfer of training of the behavioral skills to the cockpit" (Prince & Salas 1993, p. 355).

Measurement of reactions to training is easiest to implement. Most programs include a brief questionnaire polling participants about how important or useful they judge the different elements of the training to be. This kind of feedback may be useful for adjusting program contents, but does not suffice for showing program effectiveness. For example, Butler (1993, p. 249) noted that following poorly designed LOFT scenarios, which "were nothing more than a full mission simulation with a problem requiring the use of an abnormal checklist, presenting no challenge, no decisions," the crews would still claim that they "got a lot out of the LOFT that they would take back to the line." The other measures are also not too difficult to apply. Conceptual learning can be demonstrated by follow-up quizzes, and transfer of training can be demonstrated by LOFT or line checks to see if crews apply the behavioral skills in the cockpit.

To measure attitude change, one needs to develop an instrument that measures relevant attitudes and then collect before and after ratings. For example, the Cockpit Management Attitudes Questionnaire (CMAQ) was developed for this purpose. It was developed for CRM research by Robert Helmreich (1984) at the University of Texas. The rationale forwarded for measuring attitudes is that CRM training is supposed to affect attitudes (Helmreich & Foushee, 1993. p. 34).

Because attitudes have been shown to be predictors of crewmember behavior in the cockpit (Helmreich et al., 1986), changes in measured attitudes on the CMAQ are assumed to impact behaviors that cause or are linked to accidents. A more direct approach to show program effectiveness is to compare the performance of crews with CRM training to those without (Helmreich et al., 1990). Of course, this requires some instrument for evaluating crew performance. For this purpose, the Line/LOS checklist was developed (Helmreich et al., 1990, p. 577; Law, 1993). This form can be used by specially trained check airmen to rate crew performance in regular line service or in LOFT simulation (Helmreich et al., 1991). Much of the University of Texas assessment project involves using the CMAQ and the Line/LOS checklist (a.k.a. Line/LOFT worksheet) to accumulate a database of attitude and performance assessments both across organizations and longitudinally tracking individuals over time (Gregorich & Wilhelm, 1993, p. 189). Studies based on these data have produced the following results:

Training. On surveys administered at the conclusion of CRM training programs at several organizations, the majority of participants rated CRM training as at least "somewhat useful," which was the mid-point on a five point Likert-type scale (Helmreich & Wilhelm, 1991). Trainees' assessment of the actual usefulness of training may tend to be more positive than an expert's assessment (Butler, 1993, p. 249). Unfortunately, there are no data to compare how any other type of training might have been rated. Thus, the Helmreich and Wilhelm (1991) data do not provide strong support for the actual usefulness or effectiveness of the training. A similar study of the usefulness of CRM LOFT training yielded average participant ratings ranging from 5.6 for one organization to 6.4 for another using a seven point Likert-type scale ranging from *strongly disagree* to *strongly agree* (Wilhelm, 1991). Both studies indicated significant differences in ratings among organizations and among crew positions. However, such results might be caused by differences in instruction techniques, trainer effectiveness, or trainee/team cooperation and aptitude levels rather than the effectiveness of the curriculum per se.

Attitudes. A comparison of CRM related attitudes, as reflected in the CMAQ, before and after training, indicates that "attitudes show significant positive shift on the three scales of the CMAQ, communications and coordination, command responsibility, and recognition of stressor effects" (Helmreich & Wilhelm, 1991). It is not clear whether the CMAQ was administered to the same individuals immediately before and after training, or whether the data reflect average pre- and post-training scores of all individuals in each organization. In the first case, one might be concerned with re-testing effects, and in the second case, one might be concerned about changes in the makeup of the groups taking the CMAQ before and after training. The sample size of this study is not indicated, but is probably large, as the data came from a database containing information from as many as 20,000 individuals (Helmreich & Foushee, 1993, p. 32), therefore statistical significance alone is insufficient to establish the level of practical meaningfulness. Average increases at the different organizations appear to range from a little more than one point to about three points on a scale that ranges from 11 to 55. All of the reported mean baseline ratings were above 45, and none of the post training mean ratings exceeded 50. The practical importance of such a small increase is debatable. Nonetheless, these data do support the claim that CRM training does not, on average, make attitudes worse. Of concern, however, is the finding that about 5 to 15% of participants "show attitude change in a direction opposite to that intended" (Helmreich & Foushee, 1993, p. 38; cf. Helmreich & Wilhelm, 1989). Furthermore, "the data indicate that those pilots reacting most negatively are low in both task and relationships on the [Blake-Mouton managerial] grid . . . [and are] judged to be the most in need of improvement" (Cook, 1990, p. 31). It is conceivable that any benefit resulting from improved overall attitudes may be offset by the possible detriment resulting from the negative reactions of those few trainees.

Behaviors. A longitudinal comparison of one organization's average ratings of CRM behaviors in line operations using the Line/LOS checklist indicates continued improvement in the ratings obtained before implementation of CRM/LOFT training and at three, one-year intervals thereaf-

ter (Helmreich & Foushee, 1993, p. 34, Helmreich et al., 1990). Helmreich and Foushee interpret this trend as indicating shifts in organizational norms, wider acceptance of CRM concepts, and growing peer "pressure on crewmembers to conform to new standards of behavior," although this trend is confounded with other changes that may have occurred in this organization during the study period, such as, perhaps, increasing experience levels of crewmembers in the study or expansion and phase-in of the training program (Helmreich & Foushee, 1993, p. 34). It still remains to be demonstrated that performance ratings based on the Line/LOS checklist actually predict safer flight operations. Since the Line/LOS checklist was developed in tandem with the corresponding attitude-based CRM program, and the items reflect the behavioral skills and attitudes taught in the course (Helmreich et al., 1991), results based on it could be interpreted as implying little more than successful transfer of training of the course content. However, there is one study of B-52 bomber crews that found "a high correlation between resource management behavior as measured by the Line/MOST worksheet [Air Force version of Line/LOS checklist] score and both the mission performance rank and the bombing rank of the crews" (Povenmire et al., 1989, cited in Prince & Salas, 1993, p. 352). Unfortunately, there are important differences between Helmreich's attitude-oriented CRM program and the Air Force's KSA-oriented ACT program (Helmreich, 1987, Prince, & Salas, 1993).

In combination, these studies have been used to support the claim that CRM training is effective. It is important to note that all these studies lacked randomization or experimental control. None of these studies compared CRM training with any alternative training forms. A weakness of before-and-after studies is that improvements in attitude and performance can be caused by other variables, such as an organization's increased commitment and allocation of attention and resources to training, and possibly increases in classroom and simulator time for training (Helmreich & Foushee, 1993, p. 36). For example, it is possible that the factors that moved the organization to begin CRM training also caused the organization to undergo other changes which may also have contributed to the improved performance. It has been noted that organization-wide commitment to CRM concepts and standardization of all manuals, procedures, and training on CRM concepts, as well as provision of "check airmen and instructors [who act as] primary role models and agents of reinforcement" are important to the success of CRM (Helmreich & Foushee, 1993, p. 36). A fairer evaluation of the merits of CRM concepts would require keeping conditions equal for the CRM and non-CRM groups.

State of Affairs of CRM Evaluation

A number of authors have pointed out areas where further research to evaluate CRM effectiveness is needed. Their comments will be presented in chronological order.

Helmreich and co-authors (1986, p. 1198) comment: "Despite the commitment of substantial resources to providing this [CRM] training, valid formal evaluations of its effectiveness are lacking, although anecdotal reports of changes in attitude and behavior are abundant. The absence of methodologically sound evaluations of cockpit resource management training can be traced to a number of sensitive issues including federal regulatory policy, protection of individuals, potential liability, and labor-management relations."

By 1988, this state of affairs does not seem to have changed much, as Prince et al. (1988, p. 5) report: "CRM training programs have not been validated; little evaluation of program elements has been done." The working group on the effectiveness of CRM training at the NASA/MAC workshop, chaired by Captain J. Butler, could only conclude that there is an "intuitive feeling that CRM training may be effective" (Orlady & Foushee, 1987). Reasons for the "paucity of data" include all of the issues cited by Helmreich et al. in 1986 and also the "considerable difficulty associated with measurement of complex human behavior" (Prince et al., 1988, p. 5). Similarly, Hackman and Helmreich (1988, p. 290) note that "the critical areas of judgment, leadership, and decision-making are still rated subjectively."

Gregorich, Helmreich, and Wilhelm (1990, p. 689) report that: "Direct causal linkages from cockpit-management attitudes to crew process and line performance have not yet been established.

Data pertaining to this issue are sparse because of the difficulty of obtaining performance evaluations and crew-process measures, but some support for a connection has recently emerged (Helmreich et al., 1990). An experimental evaluation of the effects of crewmember attitudes on crew process and performance outcomes might best be addressed in a full-mission, high-fidelity, flight-simulation study. This approach allows for manipulation of input variables." The Helmreich et al. (1990) study was a preliminary report of the 3-year study discussed above. The suggested "experimental evaluation" apparently has not yet been done.

Most recently it was noted that "strategies to evaluate whether these [CRM] programs fully realize their goals are not yet mature" (Gregorich & Wilhelm, 1993, p. 173). Ironically, in the same volume, Helmreich (1993, p. 2) concluded that: "we are no longer concerned with the basic question of whether there is a measurable impact of human factors training but rather with a new set of questions that have arisen from earlier investigations [of CRM]."

Evaluation of CRM Program Components and Iterative Improvement

If we accept that CRM programs have positive effects on crew attitudes, performance, or skill level, then the next level of assessment is to examine the various components of CRM to try to discover which elements are most important and cost-effective, and which elements can be deleted. This may involve comparing different versions or implementation levels of CRM training programs and assessing the relative value of different aspects of CRM.

For example, Helmreich and Foushee (1993, p. 36) discovered that "several organizations in which flight operations management made a concerted effort to communicate the nature of CRM training and the organization's dedication have noted significant improvement in cockpit management attitudes even before formal training was instituted." By contrast, Helmreich and Foushee (1993) describe another organization in which both CRM and LOFT were provided, but: management support was weak, no formal human factors training was provided to new check airmen and instructors, there was high turnover among training and checking personnel, and there was limited effort to revise and update LOFT scenarios. A comparison between first year and second year post-training attitudes regarding the value of CRM and LOFT showed a significant drop. This would seem to suggest that organizational and other factors besides formal CRM training and LOFT have an effect on crew attitudes, and perhaps likewise performance.

A re-examination of the role of LOFT in CRM training could lead to savings and, possibly, even better training. Prince et al. (1992, p. 22) note that "there is no conclusive data on the effectiveness of LOFT for changing aircrew behavior," although it is "accepted by crewmembers" as a useful training experience. As previously mentioned, participants' estimation of usefulness may not be accurate (Butler, 1993). Helmreich and Foushee (1993, p. 39) list optimal LOFT use as an open issue for research, and comment: "we need to know much more about the effectiveness of this type of training in less elaborate facilities ... more limited simulations that do not encompass a full mission (SPOT: Special Purpose Operational Training) in communicating and reinforcing CRM concepts" (Helmreich & Foushee, 1993, p. 4).

A number of different CRM programs are in use at various commercial airlines, and in the various flying divisions of the military (Orlady & Foushee, 1987; Prince et al., 1993). These programs use different sets of CRM dimensions and skills, and target different sets of behaviors. There are also variations in the implementation and execution of the three program phases of awareness, practice and feedback, and continuous reinforcement. Yet, there seem to be no studies comparing the relative merits and effectiveness of the different programs. Even during the implementation and formulation of new CRM programs, no such comparisons are made. Typically, the program is designed based on reviews of literature, existing programs, analyses of the organization's accident and error reports, observations from flight simulator sessions, and surveys and interviews with experienced pilots, checkmen, and instructors (Gregorich & Wilhelm, 1993, p. 190; Prince et al., 1992, p. 18). Some organizations may adapt existing CRM training programs by simply inserting

additional "modules" dealing with their specific needs (Wiener, 1993, p. 202), or they "develop a program from the beginning" (Prince et al., 1988, p. 4). However, no preliminary testing of proposed programs or program elements or comparison of alternate program formulations at "test sites" seems to have been done. If the initial formulation of a program is inadequate or becomes outdated by changes in technology or regulations, specific features of the program may be modified (Butler, 1993; Helmreich & Foushee, 1993, p. 39; Wiener, 1993). However, procedures and methodologies for comparing the newly modified program with its predecessor are not specified.

Performance Evaluation Methodology

One of the challenges to the evaluation of CRM training programs is the difficulty of reliably measuring outcome variables, such as crew performance and behavior. One approach has been to measure attitudes, instead of actual performance, using the CMAQ (Helmreich, 1984). The CMAQ is a self-rating, self-report measure with 25 items. Although items are intended to assess attitudes, many of the items directly question the appropriateness of specific behaviors. For example, many of the items are of the form: "1. Crewmembers should avoid disagreeing with others because conflicts create tension and reduce crew effectiveness," or "2. Crewmembers should feel obligated to mention their own psychological stress or physical problems to other flightcrew personnel before or during a flight." These two items are directly related to the behaviors of expressing disagreements, or suggesting alternate solutions, and communicating one's psychological and physical condition to crewmembers. These attitude questions seem to have "correct" answers that individuals would recall following a CRM training session. Thus, it seems that the CMAQ might actually reflect learned knowledge, rather than changed attitudes. The CMAQ, then, might actually discriminate between individuals who are motivated to provide the "correct" answer and those who answer based on a re-assessment of long-standing attitudes and behaviors.

To measure crew behavior more directly, the Line/LOS checklist was developed (Helmreich et al., 1990, p. 577; Law, 1993). It consists of a list of eight "crew effectiveness markers," which can be rated on a 5-point poor-to-excellent scale, plus an overall technical proficiency and an overall crew effectiveness rating. The eight crew effectiveness markers are: (a) briefings, (b) inquiry/assertion/advocacy, (c) crew self-critique (decisions and actions), (d) communications/decisions, (e) leadership-followership/concern for tasks, (f) interpersonal relationships/group climate, (g) preparation/planning/vigilance, and (h) workload distribution/distractions avoided (Helmreich, Wilhelm, Kello, Taggart & Butler, 1991). Raters receive special training to learn how to reliably judge crew effectiveness. Following "training and calibration", these raters are said to "display a high level of agreement or reliability . . . in the evaluation of resource management during both line checks and LOFT" (Helmreich 1987, p. 15). However, using the form alone, without appropriate training, does not yield reliable ratings (Butler, 1993, p. 255). Part of the training involves practice rating using "a set of 52 specific concrete behavioral 'markers'" (Jones, 1993), but during actual rating, only the eight markers are used. These specific behaviors are grouped into three general behavioral clusters. For example, the "communications processes and decision behavior cluster" includes the crew effectiveness markers one through four, and provides 17 specific concrete behavioral "markers", such as "1. establishes team concept and environment for open/interactive communications," and "2. identifies potential problems such as weather, delays, and abnormal system operations. Sets expectations for how deviations from S.O.P. are to be handled." The use of the specific concrete behavioral "markers" increases objectivity, because there is more certainty in judging the existence of such specific behaviors than rating broad behaviors. The difficulty of remembering these 52 concrete behavioral markers, however, may be a reason why current and repeated training is so important for the evaluator. It is not clear whether data have been collected to establish parallel forms reliability between the 52-marker practice form and the 8-marker line/LOS checklist.

The program developed for the military by Prince et al. (1988, p. 6) included "the development of measures of effectiveness and an evaluation plan" as an

integral part of the program from the outset. Calling it "the specific-behavior approach to training" (Prince et al., 1992, p. 26), they focused their efforts on identifying "particular behaviors within skill dimensions that result in effective aircrew performance, and thus, require training. . . . In order to train and evaluate crew coordination, the construct of coordination is distilled into knowledge, skills and abilities (KSAs) and expressed as a set of specific targeted behaviors." Recall that their skill dimensions were "(1) communication; (2) situational awareness; (3) decision-making; (4) leadership; (5) mission analysis; (6) adaptability/flexibility; and (7) assertiveness" (Prince & Salas 1993, p. 358; cf. Härtel, Smith & Prince, 1991; Prince et al., 1988, p. 11). Each of these dimensions has a set of specific targeted behaviors, that is used both for training and evaluation. To illustrate this, the specific skills for decision-making are: "cross-checking information sources, anticipate consequences of decisions, use data to generate alternatives, gather pertinent data to generate alternatives, evaluate information and assess resources, identify alternatives and contingencies, and provide rationale for decisions" (Prince & Salas, 1993, p. 358; cf. Härtel, 1991, 1992; Härtel, Smith & Prince, 1991).

Fowlkes, et al. (1992, p. 342) present a method for a more objective performance rating system: "The methodology is a form of structured observation in which (1) opportunities for crews to demonstrate aircrew coordination skills in training scenarios are provided utilizing carefully structured scenario events, (2) acceptable crew responses to each of the events are determined a priori by utilizing NATOPS [Naval Air Training and Operating Procedures Standardization], SOPs and subject matter experts (i.e., military aircrews), and (3) appropriate responses to events are scored as either present or absent." A separate rating form, listing all events and all acceptable responses for each event, would have to be created specifically for each scenario. These acceptable responses are named TARGETs (Targeted Acceptable Responses to Generated Events or Tasks). This approach forces scenarios to be carefully scripted and followed. Fowlkes et al. (1992, p. 346) explain: "TARGETs are behaviors that the instructor can plainly observe and determine as being present or absent. This functions to simplify the task for instructors" and should result in very good inter-rater reliability. A small study resulted in an r = 0.94, indicating a high level of agreement between observers" with "minimal instructor training" (Fowlkes et al., 1992, p. 347). It is not clear if such an instrument could be developed for use in line checks or if the a priori specification of TARGETs constrains the kinds of problems that can be included in a scenario.

Common errors. A shortcoming of all of these evaluation methods is that they do not directly measure performance outcome variables, such as flight safety or efficiency. Instead, they are specifically created to measure CRM related knowledge, skills, behaviors, and attitudes. Furthermore, in an effort to continually improve the CRM training program, the curriculum is repeatedly reassessed and incrementally refined (e.g., see Figure 3 on p. 9 of this document, or Figure 6.4. of Gregorich & Wilhelm, 1993, p. 190). Likewise, the assessment tools are periodically revised to better measure CRM training outcomes (Gregorich, Helmreich, & Wilhelm, 1990, p. 686). The danger exists that, if the outcome assessment tools are not independent of the training program, which seems to be the case in much of CRM practice, then, as both are repeatedly refined in tandem, the assessment instruments become "tuned" to the training program but can diverge from actual outcomes that have direct consequences, such as safety, efficiency, and mission effectiveness (Helmreich & Foushee, 1993, p. 40).

The low frequency of actual flight mishaps makes infeasible their use as an outcome performance measure. However, incidents, errors, deviations from SOPs and other "unwanted occurrences" that do not result in reportable accidents or mishaps, occur more frequently (Cook, 1990, p. 33; Helmreich & Foushee, 1993, p. 5; Jones, 1993; Orlady, 1993, p. 452). One source for information about unwanted occurrences is the ASRS (Aviation Safety Reporting System). The ASRS is a system whereby flightcrew members can report such occurrences anonymously. Orlady (1993, p. 452) remarks, "Virtually every report made to the ASRS is an 'unwanted occurrence,' and virtually every 'unwanted occurrence' represents a breakdown or a weakness in the aviation system."

Collection and analyses of such reports could be used as supplemental outcome assessment tools of crew training programs, if this could be managed without violating their confidentiality. Jones (1993) recommends that the conditions leading up to accidents be examined and used to create a database "with records of the presence or absence of selected behaviors . . . in crew-related mishaps." It "would be best if . . . successes as well as failures" were included. "A baseline measure for the real behavior of flightcrews in their normal environments in the absence of reportable incidents could serve as a reference for the extent of deviation from standard behavior when incidents do occur" (Jones, 1993). If this database also contained information about crew training and other input variables, then some inferences could be drawn about causal factors of accidents and also of accident avoidance. Anecdotal evidence of the benefit of crew training in minimizing consequences of serious mishaps already exists (Helmreich et al., 1987).

Another avenue for collecting more objective performance data is suggested by a study (Chidester et al., 1990) which used "technical errors . . . coded from computer records and videotapes" as a measure of performance (Helmreich & Foushee, 1993, p. 10). This approach could be used in combination with experimentally controlled simulator studies (e.g., Povenmire et al., 1989, cited in Prince & Salas, 1993; Thackray & Touchstone, 1983; Wiener, 1993, p. 208).

SUMMARY

CRM training originated to reduce human errors in flight operations. Its goal continues to be teaching effective utilization of the resources available to the aircrew. Typically the main focus is on personal and interpersonal skills development. CRM training programs usually include three components: CRM awareness lessons, practice and feedback sessions, and recurrent refresher courses. High fidelity LOFT simulations are generally recommended for practice and feedback sessions. Efforts to assess program effectiveness are usually limited to self-report questionnaires and, in some cases, ratings using special checklists. We were unable to identify any studies using experimental controls, randomized samples, control groups, or comparisons to any other training alternatives to support the effectiveness of any particular formulation of CRM training. Lack of such scientific evidence, however, does not imply that the CRM programs in use are ineffective. It simply means that their evidence for effectiveness is derived subjectively from the expertise of the professionals involved in creating the interventions, rather than from experimentation, or other objective evidence.

Controller Resource Management

The idea of applying cockpit and crew resource management (CRM) concepts to air traffic control specialists (ATCSs) is not new. It seems to have arisen from the recognition that air traffic controllers are part of the aircrew's external environment (Orlady, 1993, p. 464; Sherman & Helmreich, 1993, p. 2) and that the aircrew's need to communicate with ATCSs makes each a resource of the other (Helmreich & Foushee, 1993, p. 24). As CRM evolved from cockpit resource management to crew resource management, the role of others in the aviation system received increasing attention (Helmreich & Foushee, 1993, p. 3). Since the ultimate goal of CRM is to improve flight safety and "controllers play an important role in successful, safe air transport" (Sherman & Helmreich, 1993), a natural extension of CRM was to include controllers in this safety improvement effort. The role of ATCSs has also played a part in line-oriented flight simulation (LOS) for aircrews, since realistic air traffic control and pilot-ATCS communications are an important part of the experimental or training environment (Helmreich et al., 1993, p. 493). However, the ATCS's communication in LOS is typically "scripted and performed by the experimental confederate" (Helmreich et al., 1993, p. 493), so the results of such research have little impact on ATC. Compared to cockpit resource management, research into issues of air traffic crew resource management (ATCRM) is relatively scarce and is just in its infancy (Helmreich et al., 1993, p. 493; Sherman & Helmreich, 1993). Nonetheless, some research findings related to human factors in ATC are relevant to ATCRM.

The need for ATCS human factors training is supported by analyses of incident and accident reports. Several accidents have been blamed on "ATCS's lack of adequate communication and coordination with air crews" (Sherman & Helmreich, 1993, p. 2). Furthermore, "human error was listed as the primary cause of most operational errors and deviations," according to the FAA Profile of Operational Errors in National Airspace System Report for 1989 (ATTE, 1992, p. 2). Finally, the growth and evolution of the airline industry has placed new demands on ATC, requiring more emphasis on workload management and team skills (ATTE, 1992).

Implementations of ATCRM

The specific composition of the team and the specific nature of the team's task all affect the team's resource management needs. Some organizations have adjusted their CRM training programs by simply inserting additional "modules" dealing with their specific needs (Wiener, 1993, p. 202), while others have found it necessary to "develop a program from the beginning" (Prince et al., 1988, p. 4). Despite all the differences in the composition of flightcrews, arguably they have more in common with each other than they do with air traffic control teams. Therefore, it is probably not advisable to attempt a straightforward application of FDCRM to ATCRM.

According to the FAA Office of Air Traffic Program Management, implementation of CRM-like programs in ATC have already begun at a number of sites across the country. Information on some of these programs was not available and thus, the following comments may not apply to all sites that have CRM-style training programs. The following provides some of the history and current work on applying CRM programs to ATC.

The ARTCC facility in Seattle implemented their "Controller Awareness and Resource Training" (CART) program in 1988. Other centers, including Boston and Memphis, for example, have followed this lead and adopted similar programs. The emphasis of the CART program is on teamwork, communication, awareness, and the effect of human factors on air traffic control. A conference was held in 1991 to facilitate and coordinate further growth of ATCRM training. An outcome of this meeting was the creation of the Air Traffic Teamwork Enhancement (ATTE) Steering Committee in 1992. Its mission is to develop CRM training materials and to help other facilities to implement similar CRM programs.

The ATCRM conference represented an important first step in making CRM training part of standard ATCS training. One of the goals of the conference was to "define the issues and/or problems related to CRM in the operational environment." (ATCRM conference, 1991, p. 2). Four main issues/problems were identified by most of the workgroups at the 1991 conference:

> **Work Environment**: The current work environment promotes the creation of individuals rather than teams. Controllers are thrust into an environment that requires them to use their energy to survive rather than to grow and learn. The environment is supported by the following elements: poor communication, attitudes, and negative reinforcers.
> **Communication**: Poor communication at all levels of the organization causes conflict and disagreement rather than effective relationships . . . [and] prevents a group's ability to develop a common goal and work as a team.
> **Attitudes**: Current attitudes (e.g., macho controller, can't ask for or accept help) need to change. These attitudes also prevent teams from forming and working together effectively.
> **Negative Reinforcers**: The current environment uses and promotes a punitive system whereby people are swiftly punished for making mistakes yet not trained how to avoid making those mistakes. A system should be developed that recognizes and rewards positive performance and promotes a train to succeed philosophy . . . [The en route workgroup also included the issue of trust among ATCSs.]
> **Trust**: Increased trust is needed among team members, between controllers and pilots, between controllers and supervisors/managers, between controllers working other positions, and in the system. Ensure technical competence of all members. Provide mentorships to help incorporate new members. (ATCRM conference, 1991, p. 3).

Individuals may hesitate to fully utilize resources that they cannot trust to be reliable, competent, or cooperative. A key outcome of the conference was the recommendation to establish a central steering committee to develop and guide CRM training to address these issues. The course materials that were developed by the resulting body, the ATTE Steering Committee, provide some insight on the implementation of CRM training in the ATC community.

Air Traffic Teamwork Enhancement Program

The stated goals of the ATTE program are (a) "to encourage controllers to work as team members rather than as solo performers," (b) "to improve controllers' abilities to manage their personal and team resources more effectively," (c) "to enhance controllers' abilities to relate to one another," (d) "to improve controllers' team communication and problem-solving skills" (ATTE, 1992, p. 2).

A three-day workshop designed to address these issues is described in the ATTE controller workshop manual (ATTE, 1992). The workshop consists of five sections: "1: understanding air traffic teamwork, 2: communicating with others, 3: being a resource, 4: managing conflicts, and 5: managing stress." The ATTE workshops are designed to include both supervisors and controllers in sessions, and are conducted by a facilitator who has "technical credibility and is operationally current," is "respected by management and the work force" and possesses "good listening and communication skills" (ATTE, 1992, p. 5). The typical workshop lasts three full days, has class sizes of twelve to fifteen participants, and uses two facilitators (ATTE, 1992, p. 10). At the end of the workshop, students are asked to indicate the extent of their knowledge of nine general areas, such as "awareness of all potential resources," "identification of obstacles to using resources and strategies for overcoming the obstacles," and "listening to others" as it was before training and now after training (ATTE, 1992, p. 20). The evaluation form also requests participants to provide an overall rating of the workshop, a description of what they plan to do differently following the workshop, what the most effective parts of the workshop were, and suggestions to improve it.

The facilitator never lectures; instead, videotaped presentations, group discussions, team presentations, analyses of case studies, team exercises and activities, and self-assessment questionnaires and critiques of videotaped behaviors are used to convey course content. To illustrate, one of the team exercises involves using straws and pins to design and construct a tower that is tall, stable, and aesthetically appealing. Teams of five or six participants are formed and one leader is chosen for each team. At the end of the time period allotted for designing the tower, the team leaders are informed that they, and their tower designs, have been reassigned to new teams. This is expected to cause some interesting team and leadership dynamics because of the sudden change in venue and because, in each team, only the team leader was involved in designing the tower that the team must now construct. These are discussed and related to work situations.

One of the essential elements of CRM training is program evaluation. The evaluation form used in the ATTE workshop (ATTE, 1992, p. 20) is mainly useful for course feedback and improvement. Availability of a standardized, reliable, and valid instrument to measure training outcomes makes it possible to compare the effectiveness of different training programs or program elements with each other, and to compare training effectiveness at different facilities and positions (Sherman & Helmreich, 1993, p. 2). For this purpose, "the CMAQ [was revised] for use with the en route controller population." The resulting questionnaire, named CRMAQ, consists of 25 statements reflecting controller resource management concepts that are rated on a 5-point scale, ranging from "disagree strongly" to "agree strongly" (Sherman & Helmreich, 1993, p. 19). Some sample items include "1. Crewmembers should avoid disagreeing with others because conflicts create tension and reduce crew effectiveness" and "4. 'R'-sides should not dictate technique to their 'D'-side." The CRMAQ was given to 390 air traffic controllers to see if "the factor structure of the CRMAQ generalizes across different

controllers and facilities" and to see if it "taps the same factors as the" CMAQ (Sherman & Helmreich, 1993, p. 13). "Four attitude scales were obtained. . . . Three of the scales corresponded roughly to the subscales found in the CMAQ. . . . Consistent differences were found between facilities and between different controller job descriptions. . . . Due to its short length, it [CRMAQ] is not a highly reliable scale." Also, several of the items were found to be inadequate and needed to be rewritten. Sherman and Helmreich indicated that a revised version of the CRMAQ would be forthcoming (1993, p. 16). Since the CRMAQ is an adaptation of the CMAQ, any criticisms or weaknesses of the CMAQ may also apply to the CRMAQ.

Several significant steps have thus been taken toward implementing CRM training in the Air Traffic Control system. First, the need for CRM training was identified. Second, existing programs both within the ATC community and in the aircrew community were reviewed. Third, a conference was held to bring together subject-matter experts to draw up specifications and plans for the program. Fourth, a steering committee was established to develop training materials and provide assistance to potential program administrators. And, fifth, development of evaluation and research tools has begun. So far, the main contributions of flightcrew resource management (FDCRM) to ATCRM have been the adoption of CRM concepts, training materials and methods, evaluation tools, and program implementation strategies.

Transition from flight deck to ATC

Further contributions to the development of ATCRM can be found in the development of FDCRM. It is important to keep in mind that CRM concepts were initially developed in response to specific problems in flightcrew operation, and are based on both theoretical and empirical research specifically addressing flightcrew issues. In contrast, ATCRM is an adaptation of an already existing body of CRM concepts and methods. It is important that the development and formulation of ATCRM concepts be based on research specifically identifying and addressing controller issues. The ATCS profession has its own unique challenges and needs, and should not be treated as an extension or variation of flightcrews. Arguments, such as, "although there are differences between the tasks of pilots and controllers, the environment is similar in basic ways" (Sherman & Helmreich 1993, p. 7), tend to gloss over the significant differences in the human factors and team aspects of the ATCS and flightcrew.

Differences among flightcrews with respect to such factors as mission, team makeup, experience level and training, cockpit layout, and instrumentation are all considered to be important to FDCRM (cf. Helmreich & Foushee, 1993, p. 39; Leedom, 1990; Orlady, 1993, p. 455; Prince & Salas, 1993, pp. 341-344; Wiener, 1993, pp. 201, 213). For example, Leedom (1990) concludes that "differences in the type of flying required by the Army and the civilian aviation communities, makes implementation of much that is taught in current CRM programs questionable" (cited in Prince & Salas, 1993, p. 343). Despite the differences among flightcrews, arguably, they have more in common with each other than they do with ATCSs. Therefore, it is likely that substantial revision of existing FDCRM programs, or development of a custom-tailored program "from the beginning" may be necessary to meet the needs of the ATCS profession (Prince et al., 1988, p. 4). Nonetheless, the theories, methods, and approaches used in the development of FDCRM programs can be used as a guide for developing ATCRM programs, as Helmreich (1993, p. 493) points out, "the same analytic strategy employed to identify problems [in CRM] is applicable in the ATC setting."

What is an ATC team?

First, it is important to analyze the structure of the ATCS team and task and to identify problems and weaknesses related to human factors and resource management. Recall that a team was defined to be "a small social system in which (a) membership is clearly defined (i.e., one can readily distinguish members from nonmembers), (b) members have differentiated roles to play in pursuit of some common purpose, and (c) the team as a whole manages transactions with other individuals and groups as it goes about its work" (Hackman, 1986).

For example, in the enroute setting, a team of controllers could be defined either as the pair of D-side and R-side controllers working a given position, or as the "crew," consisting of all controllers that administratively report to the same supervisor. Either definition fits the three characteristics of a team, but each definition has some disadvantages. For example, two controllers are not always assigned to work an enroute position. Inclusion of supervisors and specialists in the team may be difficult, because they may not be members of the ATCS bargaining unit. Also, it is important to note that the configuration of ATC teams is not the same in enroute, TRACON, and tower settings. The group dynamics in some of these situations may not even warrant a team approach, as such. It may be useful to view teams on a continuum ranging from tightly interacting and interdependent members to loosely connected individuals who cooperate with each other. The question of how, or even whether, to define an ATCS team in terms of specific ATCS tasks is not at all trivial, and requires thoughtful consideration supported by solid scientific research.

An important difference between flightcrews and controllers is the length of time they stay together. Airline flightcrews typically work together for about a month and then are re-assigned (Hackman & Helmreich, 1988, p. 288). In contrast, although the R-side and D-side controllers change about every hour, in most cases, especially in towers, controllers work as part of a controller "crew" and tend to work together for a relatively long period of time, often years. Many ATCSs can get to know each other fairly well. Therefore, team formation is likely to evolve differently for controllers than for flightcrews (cf. Hackman, 1987). On the other hand, flightcrew members face the challenge of having to coordinate with a team that they may have known for less than an hour. In contrast, controllers often have to deal with issues of mature groups, such as how to deal with difficult individual team members, and how to develop trust among team members (cf. ATCRM, 1991).

ATC Team Structure

Another difference is the authority structure, or chain-of-command in the team (see Figure 4). Authority structure is important because team dynamics among members at the same authority level are very different than among members of different authority levels. In large airline jets, the typical flightcrew has three members in a completely vertical team structure, that is, the captain is at the top of the authority structure, the first officer is next, and the second officer is at the bottom. The two-member crews also have this vertical authority structure with the captain being the leader and the first officer the subordinate. The cabin crew is ranked below the flightcrew.

An enroute controller crew, on the other hand, has a much more lateral or horizontal authority structure, with one supervisor at the top, and a number of R-side and D-side controllers below at the same authority levels, respectively. Furthermore, the authority structure is not as strict with ATCSs as it is in flightcrews. For example, the D-side controller may have substantially more experience than the R-side controller he or she is working with, and the R-side/D-side designations

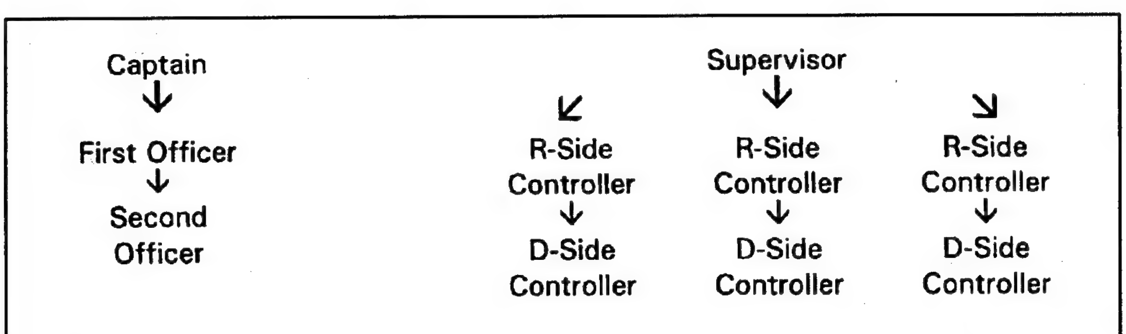

Figure 4. Simplified comparison of flightcrew and enroute controller authority structure.

are not permanent. A controller may work the R-side at one time; the D-side at another time. By comparison, the first officer is always subordinate to the captain of a flightcrew. Dynamics and communication patterns among ATCSs also differ as a function of work setting and mission. In radar settings other than enroute, for example approach control, the ATCS often works more independently of other team members, except when using an R-side/D-side pair. In the non-radar setting, however, more verbal communication takes place among all ATCSs. The question of who the members of an ATCS team should be, and what the dynamics of the team are, will require further research, but is clearly a more complicated issue than it is for flightcrews. Likely, more than one type of ATCS team will have to be defined.

ATC Team Tasks

The ATCS task is also fundamentally different from the flightcrew's task. The flightcrew, essentially operates a complex machine, the aircraft. That leads to a more singular focus of the team. Communication between R-side and D-side seems to include a lot of non-verbal communication, whereas communication with flightcrews and controllers at other locations is exclusively verbal. The focus of the controller's activity is also much more diffuse. The few points mentioned here are sufficient to indicate the need for a much closer analysis of the sub-components of the controller's task both to identify specific CRM skills and behaviors that the controller needs, and to specify the form and structure of various types of controller teams.

The general definitions of CRM apply to ATCRM equally well. ATCRM can be defined as "the management of information, cognitive work, communication, and actions needed to accomplish a time-constrained task" (Orasanu, 1990, cited in Sherman & Helmreich, 1993, p. 4), or using the SHEL terminology, ATCRM "is the effective utilization of all available resources—hardware, software, and liveware—to achieve safe, efficient flight operations" (Lauber, 1987, p. 10; cf. Helmreich & Foushee, 1993, p. 4; Lauber, 1984). Recall that S, H, E, and L refer to Software, Hardware, Environment, and Liveware. The SHEL model is useful for enumerating, and visualizing the human factors that are involved in resource management in a given situation (see Figure 5). The controller is represented by the center L box, and the L-L connection represents all communication and information flow between the controller and other humans in the system that the controller has access to, including flightcrews, flight management, and other controllers in the same sector and in other sectors. The L-H interaction is the human-machine interface.

Automation of controller functions might have important consequences on a controller's cognitive functioning (Hopkin, 1989). For example, Thackray and Touchstone (1982) argue that the introduction of highly-automated air traffic systems could change the role of the ATCS from that

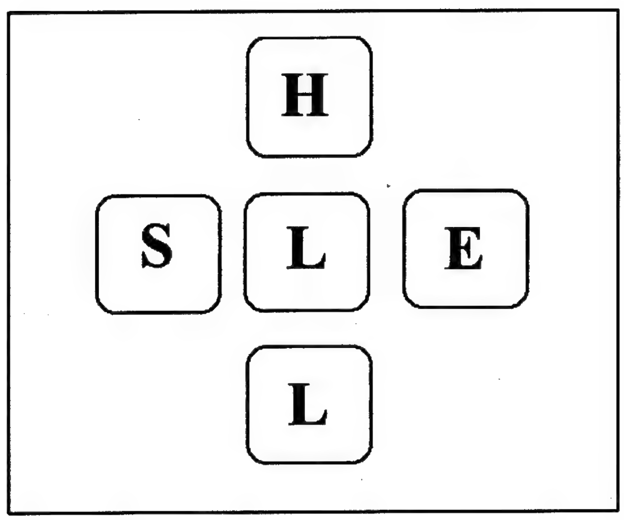

Figure 5. SHEL model.

of an "active planner" to that of a "passive responder to alternate courses of action presented by the computer." Advanced automation of ATC tasks could lead to potential problems such as "controller complacency, inattentiveness, boredom, and reduced readiness." The L-E link represents the interaction of the controller and his or her environment, and apart from a controller's environment being uncomfortable, it can have subtle, but serious influence on his or her performance. The S-L link represents the human factors involved in the non-physical aspects of the system, such as procedures, manuals, and even computer programs (Hawkins, 1993). Each of these

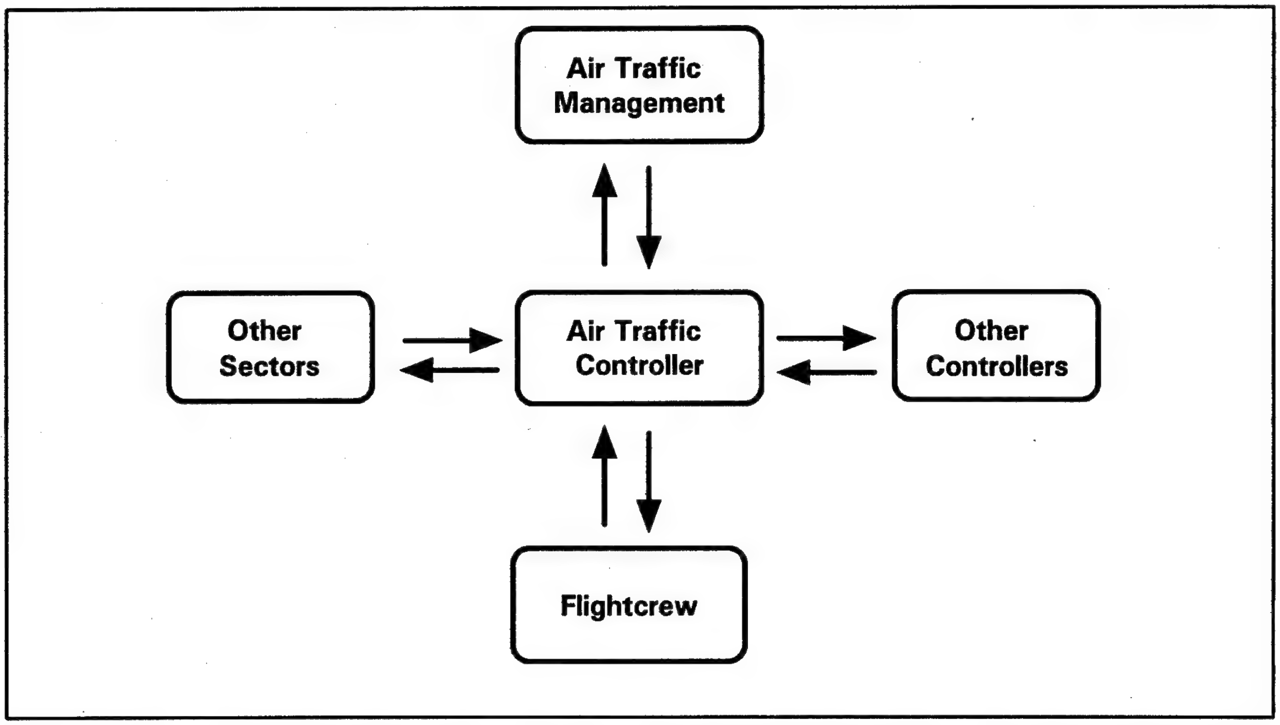

Figure 6. The role of the controller in information transfer, from Kanki & Palmer, 1993.

interactions represents important human factors and possible sources of human error (Hawkins, 1993, p. 34).

Research of ATCS involvement in mishaps helps define some criteria for ATCRM. Kanki and Palmer (1993) have studied the relationship of accidents and incidents with communication patterns within the cockpit, between the cockpit and ATC, and within ATC. One analysis of over 28,000 reports of incidents to the ASRS from 1976 to 1981 by pilots and ATCSs found that over 70% of these reports involve "the transfer of information." The reports focused on "pilot-controller interactions and controller communications more often than on within cockpit communications" (Billings & Cheaney, 1981). The most common reasons for communication failures were because "(1) the person who had the information did not think it necessary to transfer it or (2) that the information was transferred, but incorrectly" and not due to an unavailability of the needed information (Billings & Cheaney, 1981; Foushee & Manos, 1981; Kanki & Palmer, 1993). Findings like these suggest that at least some of the CRM principles should be useful for ATCSs. The ATCS plays a "pivotal role" in the information transfer or communication structure in today's aviation system (see Figure 6). Some factors that interfere with information transfer have been identified from ASRS reports to include "frequency saturation, high workload, and inadequately presented data," and behaviors include "distraction, failure to monitor, and complacency" (Kanki & Palmer, 1993). These factors can become exacerbated by technological solutions, if they are designed without sufficient regard to human capabilities (Hopkin, 1989).

ATC Team Roles

The pivotal role of the ATCS in the information transfer scheme of the aviation system (see Figure 6) underscores the importance of communication skills to the ATCS. An investigation of factors affecting the different types of communication that are part of the controller's role was conducted by Human Technology Inc. (HTI, 1991). HTI reported that speech patterns that were characteristic of ground-air communications were very different from controller-controller communications and ground-line communications (cited in Helmreich et al., 1993, p. 493). Speech pattern differences could be explained by an analysis of the tasks unique to each category of communication. Further research into the different aspects

of ATCS communication can help to clarify the specific skills and behaviors that ensure effective controller functioning. Furthermore, communication is not the only factor leading to operator errors. For example, such errors can also occur when radar and other equipment malfunctions, when operators misread the information presented on the radar scope (e.g., speed, altitude), ignore important information (e.g., aircraft capabilities, weather changes), or make poor decisions. More research is needed to analyze the other human factors involved in the ATCS task and to find corresponding ATCRM skills and interventions.

RECOMMENDATIONS

Our goal for this paper was to provide an overview of the scientific literature regarding Crew Resource Management (CRM), and also to determine what lessons had been learned that could be applied to ATCS training. Our conclusion is that direct adaptation of existing FDCRM programs to the ATCS profession is not advisable. Instead, the specific needs of the ATCS profession should be systematically analyzed, and appropriate training interventions should be created to meet unique mission and facility needs. Furthermore, systems for rigorously and objectively analyzing the effectiveness of the resulting training programs must be established from the outset, so that training can be continually improved and updated. We continue to use the designation CRM, although the resulting programs may not be similar to the CRM programs commonly in use. The following sections detail our recommendations for developing and implementing ATCRM training programs.

ATCRM Program Development

A review of the strategies used to develop FDCRM programs leads us to believe that several steps are important for the design and operation of effective CRM programs. Based upon lessons learned by others, these steps should benefit ATCRM program development. It should be stressed that ATCRM is not a futuristic concept. Programs have already been implemented at ATC facilities across the country that include significant elements of CRM. The following description provides a framework, based upon our review of the published CRM literature, that should provide a structure for guiding the evolution of existing and future training to ensure an optimally effective ATCRM program.

Needs Analysis

The first step to developing an ATCRM program should be to conduct a systematic needs analysis to formally identify and describe which ATCS behaviors contribute to accidents, errors and other "unwanted occurrences," and which behaviors lead to efficient, safe operations. Examinations of accident and incident reports, interviews with experienced controllers and supervisors, and observations of both simulated and actual operations would produce a comprehensive perspective of the ATC job.

Job Analysis

A second step should be to examine the needs analysis data to review ATCS tasks in various environmental and mission-related configurations. Analyses should help identify the resources available to the ATCSs, and determine whether these are used to the greatest benefit. These resources include software, hardware, and liveware. Poorly designed or malfunctioning hardware and software can directly or indirectly lead to accidents by placing unnecessary demands and stresses on operators, leading to unsafe operator behaviors. The specification of liveware resources should address the issue of teaming. It may become evident that some positions are not suitable for teaming at all, and in some cases, it may be difficult to define who the members of the team should be. An analysis of information flow, other interactions among individuals, and overlap of tasks can help identify potential team members. The life expectancy of a team should also be examined. If team membership is very transient, then the concept of a team may not be helpful or appropriate without changes in staffing policy. For example, although an ATCS and pilot interact and communicate during a particular segment of a flight (e.g., taxi, departure) the short duration of the relationship hardly allows for traditional concepts of team formation.

Setting Goals and Objectives

Third, the information resulting from the needs analysis should be used to develop an inventory of general skills or CRM dimensions that ATCSs will need to function efficiently and avoid errors. The purpose of this step is to clarify what the ATCRM program is supposed to achieve. A systematic comparison of effective and less effective controllers can shed light on which skills are most important for effective team operations. The final inventory should form the basis for training goals and program objectives.

Training Content Identification

Fourth, effective CRM dimensions or skills should be operationalized by enumerating the specific component behaviors and the knowledge, skills, and abilities (KSAs) that can be taught and evaluated. Definitions of these specific behaviors and KSAs will depend on the specific task layout, position of the controller, and characteristics of the facility. These specific behavior-based definitions should be periodically validated. Validation helps to ensure that ATCRM actually leads to optimum performance and does not have any negative consequences or side effects, such as reducing morale, job satisfaction, or creativity.

Evaluation of Training Effectiveness

Fifth, objective and valid measures of performance must be developed and used to evaluate the effectiveness of ATCRM training. A review of CRM evaluation methods suggests that a combination of approaches is most useful. For example, one approach might use a checklist to detail very specific behaviors, and derive a quantitative score based on the simple presence or absence of the behaviors in a subject's performance (cf. Fowlkes et al., 1992, p. 347; Chidester et al., 1990). A second approach could use trained raters to evaluate more complex qualitative aspects of performance, such as good team coordination, clear communication, and effect on team members. A combination of these two approaches would provide both objective and subjective criteria for evaluation. The purpose of these evaluations should be limited to training and program development, and steps should be taken to protect and ensure each ratee's confidentiality. If individual or team effectiveness assessments are desired for certification or similar purposes, a separate program should be developed that disassociates certification from training. Attempting to train and certify at once is not only ethically questionable, the potential chilling effect on trainee behaviors can potentially undermine the optimal effectiveness of both activities.

Program Updating

Sixth, a system of curriculum evaluation and revision must be implemented to refine the training and keep it current. This system should consist of three elements: (a) regular evaluations of the course content and classroom execution, (b) an established mechanism for course modification, and (c) pilot testing to ensure training (and any modifications) achieves desired on-the-job outcomes. Regular evaluations of the course content and execution should use performance measures and other methods for judging the effectiveness of the training efforts. Periodic re-assessment of training needs should be included whenever there is a significant change in the task design or work environment (e.g., changes in air traffic patterns, automation, or regulations). Next, a formal mechanism for course redesign will provide the means necessary to make ATCRM program changes when they are needed. Such a mechanism might consist of a committee or panel comprised of experienced instructors, controllers, managers, and education experts, who would be regularly involved with program development and redesign. Lastly, there must be a means for evaluating the impact of the ATCRM program (including any training modifications) on job performance. Both should be in place prior to full implementation. This will ensure that trainers and curriculum developers have sufficient valid information on which to base changes in techniques, curriculum content, etc. to ensure optimum effectiveness.

Program Implementation

A review of the methods and strategies used in the implementation of FDCRM training programs suggests to us that the following elements should be part of ATCRM training programs.

We mentioned earlier that most CRM training programs include three phases: *awareness*, *practice and feedback*, and *continual reinforcement*. The awareness phase introduces and explains CRM concepts..The practice and feedback phase consists of practicing the CRM skills in a realistic setting, such as a simulator or actual work setting, and receiving performance feedback via self-critique, peer-critique, and by reviewing videotapes. The continual reinforcement phase reinforces use of the CRM skills by regularly repeating elements of the other two phases. Reinforcement sessions typically include annual or semi-annual refresher seminars on CRM concepts, along with practice and feedback sessions in the simulator. The implementation of ATCRM training should attempt to satisfy the goals of these three phases, although it may not be necessary to implement separate awareness, practice, and reinforcement phases.

Awareness

ATCRM programs should familiarize trainees with the basic concepts of CRM. Awareness training should pursue the goal of creating an understanding of the role human factors plays in effective team performance. Videotapes of teams offer a useful means for illustrating effective and ineffective ways of interacting to get the job done. Of course role playing, exercises, and simulations conducted during the practice sessions also advance conceptual awareness when coupled with good feedback.

Practice and Feedback

ATCRM training must provide opportunities for practice and feedback of CRM skills and behaviors. These can include high fidelity ATC simulations, desktop computer simulations, role play, feedback from actual on-the-job critical incidents using either high fidelity systems like FAA's new SATORI system (Rogers & Duke, 1993) or low fidelity systems such as videotaping routine work station activities. The use of high fidelity ATC simulators such as TRACONPRO (Klass, 1991; Nichol, 1991) for this purpose can be valuable, because it provides control over the scenario and errors do not result in dangerous situations. In addition, use of low-cost simulators, possibly using networked personal computers may also prove to be useful. It might be possible to achieve a reasonable level of realism economically by combining standard desktop computers with such custom-designed components as special keyboards, software, and workstations which position monitors and controls in standard work station configurations (cf. Härtel, 1991, 1992, 1994a).

Another option is to develop role-playing exercises that simulate group dynamics, and problem situations that require team coordination and resource management skills to solve. The tower building exercise used by the FAA's ATTE workshop is a good example of an exercise that simulates some of the interpersonal dynamics found in some work situations. Another example is the United Airlines role play exercise mentioned earlier, in which one crewmember plays a role reacting to the scripted roles of others playing the different crewmembers (Taggart, 1993).

Of course, nothing can beat the realism of feedback produced by one's own on-the-job behaviors. The FAA SATORI system uses FAA tape logs to reproduce high fidelity radio traffic and radar images of critical incidents enabling controllers to observe first hand both ineffective and effective handling of actual situations. Videotaping of routine ATC work station operations can provide a less expensive lower fidelity feedback tool. In either case, either SATORI or videotaping produces feedback without the need for simulator equipment and support personnel. Moreover, trainees are likely to have more faith in a real-life situation than a simulation, and it presents a training mechanism that can be employed on the job which is convenient, and familiar since most ATC training is provided on the job. Of course, as in the case of formal CRM training courses, on-the-job feedback training session recordings would have to be treated with confidentiality and should be erased following the feedback session. If employees feared disciplinary action, few would participate in such a training feedback program.

Continual Reinforcement

Lastly, the continual reinforcement phase can be accomplished by establishing regular refresher sessions involving elements of the other two phases. These refresher sessions might be integrated into other regular training, or may be separate training

sessions, depending on which is more convenient for a given facility. CRM sessions should also be part of other training occasions, such as for newly-hired or transferred controllers, newly-promoted controllers (e.g., new supervisors), and for transition training (e.g., different equipment or different position).

Ideas and approaches to enhance CRM training from outside the FDCRM literature should also be investigated. One innovative example would be application of the popular notion of mentoring. The FAA has been matching less experienced controllers with veterans. By providing training for the veterans on how to teach and facilitate CRM learning for their partner, mentoring can help build teamwork into the organizational culture. Moreover, by enabling mentor pairs to work together for extended periods of time, the level of coordination and cohesiveness of controller teams can be strengthened.

Another idea is to adapt the TQM idea of quality control circles, as has been done in some FAA facilities. This might involve forming special teams of controllers and making them responsible and accountable for a certain segment of work. The team might include everyone on the same rotation, working under one area supervisor. This would yield teams of between eight and twelve members (D. Broach, FAA Civil Aeromedical Institute, personal communication, September, 1993). At some facilities, it might be advisable to split the controller crew into two or more teams to keep the team at a manageable size. Each team would hold regular meetings, possibly daily or weekly, where recurrent training and feedback would take place. These meetings could include brief refreshers of a CRM topic and discussion of any critical incidents or special situations where CRM was or would have been useful.

Team Leader Feedback. It is important that feedback be given as close to the event as possible (Salas et al., 1992, p. 20). So, for the purpose of feedback, it might be useful to have short debriefing sessions following each shift. If feedback is pursued, team leaders should receive training on how to provide feedback and which aspects of performance should be emphasized. This is because "teams maximize that aspect of performance about which feedback is given, even at the expense of other aspects of performance. Thus, team training should avoid emphasizing one characteristic of team performance at the expense of others" (Salas et al., 1992, p. 20). It may be useful to develop a feedback/debriefing form or checklist, so that all key aspects of performance are discussed, both positive and negative. Such a feedback form would need to be adjusted from time to time to maintain an even balance of all sub-tasks.

Team Feedback. Team research suggests that team members "should be allowed to develop flexible work structures," and should be encouraged to "experiment with different methods of coordinating their work flow under varying task conditions" (Chapman et al., 1959, cited in Salas et al., 1992, p. 19). Team meetings could be used for this purpose, to develop strategies and arrangements for accomplishing their duties most effectively. Furthermore, this would be an opportunity for the team to discuss and resolve any problems or team issues as they arise. However, to implement feedback sessions of this sort requires a team climate in which errors are viewed as the team's problem, not the individual's. A common concern voiced during the ATCRM conference (ATCRM, 1991) was the punitive orientation of ATCS evaluations. For example, controllers are individually disciplined for triggering OEDP (Operational Error Detection Patch) errors. By making these errors a responsibility of the whole team, the whole team works to avoid such errors and find systemic solutions, some of which might include changes in work flow management or refinements to team coordination or communication strategies. If an error is due to a problem with an individual controller, the team can provide mentoring and training, or even use peer pressure to modify a resistant controller's behavior. Team members may, at times, be more aware than the superiors of the exact nature of an individual's problem. Working in a team context would give teams more freedom to cope with such individuals.

External Feedback. The controller team interacts with a number of other groups and can benefit from a better understanding of their tasks and viewpoints. Groups such as flightcrews and other groups of controllers can help identify ways that ATC teams can improve performance. It is generally recognized that flightcrews and controllers

see very different aspects of the air traffic control job. Existing programs encourage controllers to fly along in the cockpit to observe the flightcrews' job and encourage pilots and other ATC customers to observe ATCSs in action. Facilitating regular cross-specialty observation among controllers on a regular basis could help controllers better understand their areas' effect on the others. Unfortunately, there currently is little structure to exchange feedback and allow flightcrews and ATCS teams to clarify barriers to more effective performance. Lastly, it is not uncommon to hear about status concerns between flightcrews, tower crews, flight service station crews, enroute crews, crews from high traffic facilities and low traffic facilities. Such attitudes can create barriers to effective communication and cooperation. Cross-training as well as observation might help develop an appreciation of the challenges and pressures faced by each group, and thereby improve system safety and job satisfaction for all concerned.

In conclusion, the literature we have reviewed indicates that CRM has been a widely used means for enhancing job performance and flight safety for flightcrews. CRM is beginning to emerge in other areas of the aviation field. As air traffic controllers increase their involvement with CRM, they can benefit from lessons learned from over 15 years experience with FDCRM. Hopefully, some of the information presented in this report will assist ATCRM program managers, trainers, and curriculum developers to achieve program effectiveness quickly and efficiently. Although we believe that CRM can benefit ATC, it should not be characterized as a panacea. CRM is unlikely to rectify problems caused by deficient basic skills training or inadequate safety standards. Moreover, it is unlikely that CRM will prove itself to be cost-effective in every ATC work setting and mission. We hope that the FAA will study the issues raised in this report to ensure that CRM, as it evolves, will achieve an improved quality of work life for ATCSs, improved team productivity, and improved safety for the flying public.

REFERENCES

Alderfer, C. P. (1977). Group and intragroup relations. In J. R. Hackman & J. L. Suttle (Eds.), *Improving life at work*. Santa Monica, CA: Goodyear.

ATTE Air Traffic Teamwork Enhancement Controller Workshop (1992, October 19).

Blanchard, R., personal communication, 7-25-94.

Brown, Major General D. D. (1987, May 6-8). Introduction to MAC CRM Training. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 132-144). San Francisco, California.

Bruce, Captain K. D. & Jensen, Captain D. (People Express) (1987, May 6-8). *Cockpit Resource Management Training at People Express*. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 50-55). San Francisco, California.

Butler, Captain R. E. (1987, May 6-8). Pan American World Airways flight training — A new direction flight operations resource management. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 61-67). San Francisco, California.

Butler, R. E. (1993). LOFT: Full-mission simulation as crew resource management training. In E. K. Wiener, B. G. Kanki & R. L. Helmreich, (Eds.), *Cockpit resource management* (pp. 231-260). San Diego: Academic Press.

Carroll, Captain J. E. (ret.) & Taggart, W. R. (1987, May 6-8). Cockpit resource management: A tool for improved flight safety. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 40-46). San Francisco, California.

Controller Resource Management (CRM) Conference (1991). *Overall Findings. Summary of the Meeting*. Washington, D.C.: Federal Aviation Administration.

Cook, G. N., (No date given, about 1990). *Cockpit resource management training: Are current instructional methods likely to be successful?* (Federal Aviation Administration #0389).

Cooper, G. E., White, M. D. & Lauber, J. K. (Eds.) (1979). *Resource management on the flight deck* (NASA conference publication 2120).

Davidson, Captain J. (1987, May 6-8). Introduction to Trans Australia Airlines (TAA) CRM training. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 88-89). San Francisco, California.

Federal Aviation Administration (1989). *Cockpit resource management training.* (Advisory Circular 120-51). Washington DC: Author.

Foushee, H. C. & Manos, K. L. (1981). Information transfer within the cockpit: Problems in intracockpit communications. In C. E. Billings & E. S. Cheaney (Eds.), *Information transfer problems in the aviation system* (NASA TP-1875). Moffet Field, CA: NASA-Ames Research Center, September, 1980 (NTIS No. N81-31162).

Fowlkes, J. E., Lane, N. E., Salas, E., Oser, R. L., & Prince, C. (1992). Targets for aircrew coordination training. *Proceedings of the 14th Interservice Industry Training Systems Conference* (pp. 342-352). San Antonio, TX.

Gregorich, S. E., Helmreich, R. L. & Wilhelm, J. A. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology*, 75(6), 682-690.

Gregorich, S. E., & Wilhelm, J. A. (1993). Crew resource management training assessment. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management*. San Diego: Academic Press, pp. 173-198.

Hackman, J. R. (1987). Group-level issues in the design and training of cockpit crews. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 23-39). San Francisco, California.

Hackman, J. R. (1993). Teams, leaders, and organizations: New directions for crew-oriented flight training. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 47-70.

Hackman, J. R., & Helmreich, R. L. (1988). Assessing the behavior and performance of teams in organizations: The case of air transport crews. In D. R. Peterson & D. B. Fishman (Eds.), *Assessment for decision* (pp. 283-313). London: Rutgers University Press.

Härtel, C. E. J. (1991). *Improving team-assisted diagnostic decision-making: Some training propositions and an empirical test.* Doctoral dissertation. Colorado State University, Fort Collins, CO.

Härtel, C. E. J. (1992, May). *A new decision making frame that improves team-assisted diagnosis.* Paper presented at the 7th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., Montreal, Quebec.

Härtel, C. E. J. (1994a). *A decision strategy for improving autodidactic decision making in online decision situations.* Manuscript submitted for publication.

Härtel, C. E. J. (1994b). *Why decision making sometimes doesn't work.* Manuscript submitted for publication.

Härtel, C. E. J., & Härtel, G. F. (1994). *Human decision making: How to create training interventions that do not exceed human cognitive capacity.* Manuscript submitted for publication.

Härtel, C. E. J., Smith, K., & Prince, C. (1991). *Defining aircrew coordination: Searching mishaps for meaning.* Paper presented at the Sixth International Symposium on Aviation Psychology. Columbus: Ohio State University.

Hawkins, F. H. (1993). In H. W. Orlady (Ed.), *Human factors in flight* (2nd ed.). Vermont: Ashgate Publishing.

Helmreich, R. L. (1984). Cockpit management attitudes. *Human Factors,* 26 (5), 583-589.

Helmreich, R. L. (1987). Theory underlying CRM training: Psychological issues in flight crew performance and crew coordination. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 15-22). San Francisco, California.

Helmreich, R. L. (1993). Whither CRM? Future directions in crew resource management training in the cockpit and elsewhere. *Proceedings of the Seventh International Symposium on Aviation Psychology.* Columbus: Ohio State University.

Helmreich, R. L., & Foushee, H. C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 3-45.

Helmreich, R. L., Foushee, H. C., Benson, R., & Russini, W. (1986, December). Cockpit resource management: Exploring the attitude-performance linkage. *Aviation, Space, and Environmental Medicine,* 57, 1198-1200.

Helmreich, R. L., Wiener, E. L., & Kanki, B. G. (1993). The future of crew resource management in the cockpit and elsewhere. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 479-501.

Helmreich, R. L., Wilhelm, J. A., Gregorich, S. E., & Chidester, T. R. (1990, June). Technical note: Preliminary results from the evaluation of cockpit resource management training: Performance ratings of flightcrews. *Aviation, Space, and Environmental Medicine,* 61, 576-579.

Helmreich, R. L., Wilhelm, J. A., Kello, J. E., Taggart, W. R., & Butler, R. E. (1991). *Reinforcing and evaluating crew resource management: Evaluator/LOS instructor reference manual* (NASA/University of Texas Technical Manual 90-2). Austin, Texas.

Helmreich, R. L., & Wilhelm, J. A. (1991). Outcomes of crew resource management training. *International Journal of Aviation Psychology,* 1, 287-300.

Hopkin, V. D. (1989). Implications of automation on air traffic control. In R. S. Jensen (Ed.), *Aviation psychology.* Brookfield, Vermont: Gower Publishing Company.

Human Technology (1991). *Analysis of controller communication in en route air traffic control: Report to the FAA.* McLean, VA: Federal Aviation Administration.

Jones, S. G. (1993). Human factors in incident reporting. *Proceedings of the Seventh International Symposium on Aviation Psychology.* Columbus: Ohio State University.

Kanki, B. G., & Palmer, M. T. (1993). *Communication and crew resource management.* In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 3-45.

Klass, P. J. (1991, July 8). FAA evaluating inexpensive PC-Based system for practicing controller skills. *Aviation Week and Space Technology,* 60.

Komons, N. A. (1987). *The third man: A history of the airline crew complement controversy, 1947-1981.* Washington, D.C.: Dept. of Transportation, Federal Aviation Administration, 1987.

Lauber, J. K. (1987). Cockpit resource management: Background studies and rationale (or 'and overview'). In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 5-14). San Francisco, California.

Law, J. R. (1993). Position-specific behaviors and their impact on crew performance: Implications for training. *Proceedings of the Seventh International Symposium on Aviation Psychology.* Columbus: The Ohio State University.

Leedom, D. K. (1990). *Aircrew coordination training and evaluation for Army rotary wing aircrews: Summary of research for fiscal year 1990.* Fort Rucker, AL: U.S. Army Research Institute Aviation Research and Development Activity.

Margerison, C., McGann, D., & Davies, R. (1987). Aircrew team management program. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 90 -107). San Francisco, California.

Mead, K. M. (1990, June 6). *Serious shortcomings in FAA's training program must be remedied* [Statement before the subcommittee on investigations and oversight on public works and transportation, United States House of Representatives]. Washington, DC.

Murphy, M. R. (1980). Analysis of 84 commercial aviation incidents: Implications for a resource management approach to crew training. *Annual Reliability and Maintainability Symposium* (pp. 298-306).

Nichol, R. (1991, Jul-Sep.). A new breed of simulator. *Journal of ATC,* pp. 62-63.

Orasanu, J. M. (1993). Decision-making in the cockpit. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 99-136.

Orlady, H. W. (1993). Airline pilot training today and tomorrow. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 447-477.

Prince, C., Cannon-Bowers, J., Salas, E., Owens, J., & Morgan, B. B. (1988 or later). *Aircrew coordination training: Requirements and challenges* (Federal Aviation Administration report #0488).

Prince, C., Chidester, T. R., Cannon-Bowers, J., & Bowers, C. (1992). Aircrew coordination-achieving teamwork in the cockpit. In R. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 329-353). New York: Ablex.

Prince, C., Härtel, C. E. J., & Salas, E. (April, 1993). Aeronautical decision making and consistency of crew behaviors. *Proceedings of the 7th International Symposium on Aviation Psychology,* Columbus, OH.

Prince, C. W., & Salas, E. (1989). Aircrew performance coordination and skill development. In D. E. Daniels, E. Salas, & D. M. Kotick (Eds.) *Independent research and independent exploratory development (IR/IED) programs: Annual report FY88* (Technical Report No. NTSC 89-009). Orlando, FL: Naval Training Systems Center.

Prince, C. & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management.* San Diego: Academic Press, pp. 337-365.

Rogers, M. D., and Duke, D. A. (1993). SATORI: Situation assessment through recreation of incidents. (Technical Report No. DOT/FAA/AM-93/12). Washington, DC.: FAA Office of Aviation Medicine.

Ruffell Smith, H. P. (1979, January). *A simulator study of the interaction of pilot workload with errors, vigilance, and decisions.* (NASA TM-78483). Moffet Field, CA: NASA-Ames Research Center.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Towards an understanding of team performance and training. In R. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3-29). New York: Ablex.

Sherman, P. J., & Helmreich, R. L. (1993, January 29). *The controller resource management attitudes questionnaire (CRMAQ): Review, Methodology, and Results* (NASA/UT/FAA Technical Report 93-1).

Shroyer, Captain D. H. (ret.) (1987). The Development and implementation of cockpit resource management in UAL recurrent training. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 47-49). San Francisco, California.

Taggart, W. R. (1993). Captain upgrade training: A new focus for enhanced flight operations. *Proceedings of the Seventh International Symposium on Aviation Psychology.* Columbus: The Ohio State University.

Thackray, R., I., & Touchstone, R. M. (1982). *Performance of air traffic control specialists (ATCS's) on a laboratory radar monitoring task: An exploratory study of complacency and a comparison of ATCS and non-ATCS performance.* (Office of Aviation Medicine Technical Report no. FAA-AM-82-1) Washington: Federal Aviation Administration

Thackray, R., I., & Touchstone, R. M. (1983). *Rate of initial recovery and subsequent radar monitoring performance following a simulated emergency involving STARTLE.* (Office of Aviation Medicine Technical Report no. FAA-AM-83-13) Washington: Federal Aviation Administration

Wiener, E. L. (1993). Crew coordination and training in the advanced-technology cockpit. In E. K. Wiener, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 199-230). San Diego: Academic Press.

Wiener, E. K, Kanki, B. G., & Helmreich, R. L. (Eds.) (1993). *Cockpit resource management.* New York: Academic Press.

Yamamori, Captain H. (1987). Optimum culture in the cockpit. In H. W. Orlady & H. C. Foushee, (Eds.), *Cockpit resource management training: Proceedings of a workshop sponsored by NASA Ames Research Center and the U.S. Air Force Military Airlift Command* (#N87-22635) (pp. 75-87). San Francisco, California.

☆U.S. GOVERNMENT PRINTING OFFICE: 1995 - 660-067/20002